GREEN FLOURESCENT PROTEIN INSPIRED CHROMOPHORE AS ESTROGEN RECEPTOR AGONIST-SYNTHESIS, BIOLOGICAL EVALUATIONS AND CELLULAR APPLICATION

A Dissertation
Presented to
The Academic Faculty

by

Christopher L. Walker

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Chemistry and Biochemistry

Georgia Institute of Technology
May 2019

GREEN FLUORESCENT PROTEIN INSPIRED CHROMOPHORE AS ESTROGEN RECEPTOR AGONIST-SYNTHESIS, BIOLOGICAL EVALUATIONS AND CELLULAR APPLICATION

Approved by:

Dr. Loren Williams, Advisor
School of Chemistry and Biochemistry
Georgia Institute of Technology

Dr. Terry Snell
School of Biology
Georgia Institute of Technology

Dr. Charles L. Liotta
School of Chemistry and Biochemistry
Georgia Institute of Technology

Dr. Bahareh Azizi
Kuwait Foundation for the Advancement of Sciences

Dr. Adegboyega Oyelere
School of Chemistry and Biochemistry
Georgia Institute of Technology

Date Approved: [October 01, 2018]

Dedicated to my mother and father for their endless sacrifice

ACKNOWLEDGEMENTS

No journey is successful because of the efforts of one individual but the community that supports the goal. I am grateful to the people who have supported my journey through graduate school. I'm thankful to my advisors Dr. Loren Williams and Dr. Bahareh Azizi for not only your support and mentorship but your guidance, patience and direction. I appreciate your unwavering commitment to helping me grow as a scientist and individual, you have offered your support as life has continually changed and for that I'm grateful.

In addition to my advisors the leadership and mentoring of various faculty and administrators at Georgia Tech has been outstanding with that I would like to thank Dr. Charles Liotta, Dr. Cam Tyson and Dr. Keith Oden. I appreciate the insightful conversations provided by Dr. Liotta working as a teacher's assistant under you for 3 years allowed me the chance to develop my own style as a teacher. Dr. Tyson thank you for your continually support and concern your assistance through ever transition during graduate school was appreciated. Dr. Oden thank you for your guidance, mentorship and friendship I reverence our time spent together and the lessons imparted to me during that time thank you for giving of yourself and your experiences.

In addition to the community at Georgia Tech that has supported me, my friends and family have embraced my process and journey as if it were their own. My most sincere appreciation goes to my brothers Santino, Rico and Isaac thank you for the supportive conversations, words of encouragement and belief in me this accomplishment is just as much a testament to persistence as it is to your friendship. I would also like to thank my parents who have given everything of themselves to support any dream that I could imagine since a child. As parents you both have loved and encouraged me through every phase of life, thank you for every lesson taught, for exemplifying

a hard work ethic, for demonstrating what the process looks like. There are no words that can express my true gratitude for what you have done and what you mean to me I am honored to be your son.

To my own family, my wife and son thank you for always being the reason “why”. To my wife this has been an amazing journey thank you for your support and willingness to partake in the adventure with me. To my son thank you for being everything, my heart, my laughter and joy time spent with you is the highlight of everyday and while this is a milestone accomplishment having you is the greatest accomplishment of all. Thank you for your understanding and encouragement while at the moment age does not provide you with the understanding to grasp why “dads school work” takes me away from moments as you get older hopefully you will comprehend this was about teaching you to finish what you start and early insight to what the “process” looks like.

Lastly to my grandmother Ollie May Jackson and my friend Dina Reynolds thank you for being there in the beginning of this process I only wish you could be here to see the ending. Thank you for the lasting contributions you made into my life. I miss you both every day, your love, friendship and past conversations guide me even today. Grandmother I cherish the memories and hold the moments as precious and yes in all things including this body of work “To God be the glory”.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS **v**

LIST OF TABLES **xxi**

LIST OF FIGURES **xii**

LIST OF SYMBOLS AND ABBREVIATIONS **xxv**

SUMMARY **xxix**

CHAPTER 1. Overview and Summary **1**

1.1 Motivation **1**
1.1.1 Designing of a potent estrogen receptor agonist **3**
1.1.2 Promoting selectivity within isotypes and among nuclear receptors **4**
1.1.3 Accessing the role of florescence and activation **5**
1.2 References **7**

CHAPTER 2. Green Fluorescent Protein **8**
2.1 Historical Perspective **8**
2.2 Green Flourescent Protein **9**
2.2.1 Green Flourescent Protein Discovery **11**
2.2.2 Green Flourescent Protein Structure **12**
2.3 Synthesis of GFP Chromophore **13**
2.4 GFP Photochemical Properties **15**
2.5 Fluorescent Proteins **16**
2.6 References **17**

CHAPTER 3. Nuclear Receptors **19**
3.1 Nuclear Recptors Overview **19**
3.2 Nuclear Receptor Structure (Regions and Domains) **21**
3.2.1 Activation transcription factor (AF-1) **22**
3.2.2 DNA binding domain **23**
3.2.3 Ligand binding domain **23**
3.2.4 Activation transcription factor (AF-2) **24**
3.3 Nuclear Receptor Domain Function **24**
3.3.1 Activation Transcription Facrtor-1 (AF-1) **24**
3.3.2 DNA binding domain **24**
3.3.3 Hormone response element **25**
3.3.4 Hinge **25**
3.3.5 Ligand binding domain **25**
3.3.6 Activation Transcription Factor-2 (AF-2**)** **25**
3.4 Nuclear Receptor Ligands **26**
3.4.1 Agonist **26**

3.4.2 Antagonist **28**
3.5 Mechanism of Ligand Mediated Receptors 29
3.5.1 Apo **30**
3.5.2 Halo **31**
3.6 Nuclear Receptors Role in Diseases 31
3.6.1 Estrogen Receptor **32**
3.6.2 Peroxisome Proliferator- Activated Receptor gamma **33**
3.7 Nuclear Receptors Role as Pharmacutical Targets 33
3.8 References 36

CHAPTER 4. Synthesis of GFP Chromophore Inspired Ligands as Estrogen Receptor Alpha Agonist 39
4.1 Estrogen Receptor 39
4.1.1 Estrogen Receptor Alpha **39**
4.1.2 Ligand Binding Pocket **40**
4.1.3 Ligands **41**
4.2 Estrogen Receptor Ligand Design Rationale 43
4.2.1 Hydrogen Bonding Network **46**
4.2.2 Ligand Volume and Size **47**
4.2.3 Molecular Topology **48**
4.3 Synthesis of Arylmethyleneimidazolone 50
4.4 Arylmethyleneimidazolone Chromophore Properties 54
4.5 Biological Assay Screens 56
4.5.1 Chemical Complementation **56**
4.5.2 Chemiluminescence **57**
4.5.3 Nuclear Receptors and Reporter Gene Assay **57**
4.6 Generation I Results 60
4.6.1 Chemical Complementation **61**
4.6.2 Luciferase Results **62**
4.6.2.1 Hydrogen Bonding Network **65**
4.6.2.2 Molecular Topography **67**
4.6.2.3 Ligand Volume **69**
4.7 Structure Activity Relationships Compounds 71
4.7.1 A-Ring **71**
4.7.2 Core-Scaffold **71**
4.7.3 D-Ring **72**
4.7.4 Hydrogen Bonding Network **72**
4.7.5 Ligand Volume and Size **73**
4.7.6 Molecular Topology **73**
4.8 Generation II Results 74
4.8.1 Chemical Complementation **74**
4.8.2 Luciferase Assay **75**
4.8.2.1 Hydrogen Bonding Network **75**
4.8.2.2 Ligand Volume-Size **76**
4.8.2.3 Molecular Topology **76**
4.9 Ligand Activation Rationale 80

4.9.1 Hydrogen Bonding Network 81
4.9.2 Ligand Volume (Shape/Size) 84
4.9.3 Molecular Topology-ACD ring effectiveness 86
4.10 Summary 89
4.11 Materials and Methods 90
4.12 References 94

CHAPTER 5. Investigating Selectivity Within Estrogen Receptor Beta And The Nuclear Receptor Family 98
5.1 Estrogen Reecptor Beta 99
5.1.1 Ligand Binding Pocket 99
5.1.2 Ligands 100
5.1.2.1 Selective Ligands 101
5.1.3 Physiological Role of Estrogen Receptor Beta 103
5.2 Design for Selectivity 104
5.2.1 ACD 105
5.2.2 Chirality 105
5.3 Chemical Complementation 107
5.3.1 Estrogen Receptor alpha Active Lignads 107
5.3.2 AMI Library Screen 108
5.3.3 Generation III Results(Chiral AMI Ligands) 109
5.4 Luciferin Reporter Assay Analysis fo Isotype Specificity 109
5.4.1 Estrogen Receptor alpha Active Ligands 109
5.4.2 AMI Library Screen 111
5.4.3 Generation III Results 111
5.5 Time Resolved Fluorescence Resonance Energy Transfer 112
5.6 Selective Ligand Activation Discussion Rationale 113
5.6.1 Hydrogen Bonding Network 114
5.6.2 Ligand Volume-Size and Shape 116
5.6.3 Molecular Topology 119
5.6.4 Chirality 120
5.7 Identification of Active Leads for Estrogen Receptor 120
5.8 Nuclear Receptor Selectivity 121
5.8.1 Rexinoid X Receptor Alpha 121
5.8.1.1 Ligand Binding Pocket Comparison 122
5.8.1.2 Rexinoid X Receptor Chemical Complementation Results 124
5.8.2 Retinoic Acid Receptor Gamma 125
5.8.2.1 Ligand Binding Pocket Comparison 125
5.8.2.2 Chemical Complementation Results 126
5.9 Summary 127
5.10 Materials and Methods 128
5.11 References 130

CHAPTER 6. Investigation Into Fluorescence and Binding 134
6.1 Nuclear receptor probes 134
6.1.1 Green Fluorescent Protein Fusion 135

6.1.2 Classical Receptor Tracking **136**
6.1.3 GFP Inspired Ligands as Nuclear Receptor Probes **138**
6.2 GFP Inspired Ligands as Nuclear Receptor Probes In vivo 138
6.3 GFP Inspired Ligands as Nuclear Receptor Probes In vitro 141
6.4 Summary 144
6.5 Materials and Methods 144
6.6 References 144

CHAPTER 7. Conclusion and Future Works 146
7.1 Conclusions 146
7.2 Future Works 148
7.3 References 150

APPENDIX A. Absorbance and Emission Wavelength for AMI Ligands 151

APPENDIX B. Characterization of Active Ligands 154

LIST OF TABLES

Table 4.1	Estrogen receptor α (ER α) tissue distribution and expression levels	40
Table 4.2	Arylmethyleneimidazolone synthesized ligands	52
Table 4.3	Percent yield of ligands synthesized	54
Table 4.4	Chemical Complementation Results of 1st Generation Ligands	61
Table 4.5	Exploration of the role of the A-ring hydroxyl	66
Table 4.6	Hydrogen Donor vs Acceptor Role of Ligand with His-524	67
Table 4.7	Perturbation of Ligand Core.	68
Table 4.8	AMI Molecular Volume & Fold Induction	69
Table 4.9	Isomeric AMI Molecular & Fold Induction	70
Table 4.10	Interatomic distance and fold induction	88
Table 5.1	Estrogen receptor β (ER β) chemical complementation	108
Table 5.2	Results of Estrogen receptor α active ligands screened against ERβ	110
Table 5.3	Percentage of estradiol displaced by select GFP-inspired ligands	113
Table 5.4	Molecular volume and preference for ERβ ligands and AMI ligands.	116
Table 5.5	Polar surface area of select AMI ligands and ERβ selective ligands	119

LIST OF FIGURES

Figure 1.2	Estrogen receptors ligand binding domains complexed with ligand	4
Figure 1.3	Ligand Binding Domain of Additional Nuclear Receptor	5
Figure 1.4	Turn on GFP Inspired Chromophore via ERα Binding	6
Figure 2.1	Aequorea Victoria displaying fluorescence	8
Figure 2.2	Aequorin- Coelenterazine Mechanism	11
Figure 2.3	Structure of GFP	12
Figure 2.4	Autocatalytic Formation of GFP-Chromophore	14
Figure 2.5	Color Variants of Fluorescent Proteins	17
Figure 3.1	Nuclear receptors regions-domains	22
Figure 3.2	Natural and synthetic ligands for various nuclear receptors	27
Figure 3.3	Shared Skeleton of agonist conversion to antagonist	28
Figure 3.4	Commercially available nuclear receptor ligands	36
Figure 4.1	Estrogen Receptor α complexed with estradiol	41
Figure 4.2	Known estrogen receptor ligands	42
Figure 4.3	Estrogen receptor pharmacophore	42
Figure 4.4	Rationale for development of GFP inspired AMI ligands	45
Figure 4.5	Ligand binding pocket of estrogen receptor α	46
Figure 4.6	Synthesized generation I compounds	49
Figure 4.7	Synthesis of Arylmethyleneimidazolones	50
Figure 4.8	Synthesis of tetralin aldehyde	51
Figure 4.9	Evaluating method for GFP inspired ligands against receptors	59
Figure 4.10	Generation I dose response curve in yeast	61

Figure 4.11	Chemical complementation screening of generation I assay	62
Figure 4.12	Luciferase assay generation I ligands screen	64
Figure 4.13	Structure activity relationship compounds	74
Figure 4.14	2nd Generation chemical complementation results	75
Figure 4.15	2nd generation ligands screened in luciferase assay	77
Figure 4.16	Dose Response for top ERα agonist	77
Figure 4.17	GFP chromophore inspired active ligands and fold induction	79
Figure 4.18	TFRET results of top activators and lowest activator	80
Figure 4.19	Modeling of the hydrogen bonding network with AMI ligands	82
Figure 4.20	Modeling of AMI ligands with His-524 comparison to estradiol	84
Figure 4.21	Modeling of AMI ligands in LBD pocket volume and size	86
Figure 4.22	Alignment of CW32 and estradiol in binding conformation	87
Figure 4.23	Oxygen to oxygen distance for estradiol and top two activators	89
Figure 5.1	Estrogen receptor β secondary structure with LBP inset	100
Figure 5.2	Known estrogen receptor β agonist	101
Figure 5.3	Estrogen Receptor β selective pharmacophore	102
Figure 5.4	Core structures used in estrogen receptor β ligands	103
Figure 5.5	Estrogen receptor selective and non-selective pharmacophore	105
Figure 5.6	Estrogen receptor β ligands possessing chiral centers	106
Figure 5.7	GFP inspired ligands generated to explore chirality	107
Figure 5.8	Screening of ERα active ligands against ERβ	110
Figure 5.9	TFRET dose response curve for select GFP-inspired ligands	113
Figure 5.10	Hydrogen bonding network within estrogen receptor β LBP	114
Figure 5.11	Modeling of hydrogen bonding network within ERβ	115
Figure 5.12	Modeling of volume of Estrogen receptor β LBP with ligands	117

Figure 5.13	Placement of CW32 in secondary pocket within ERβ.	118
Figure 5.14	RXRα ligand binding pocket complexed with 9-cis retinoic acid	123
Figure 5.15	RXRα binders with similar ring systems as ERα agonist	124
Figure 5.16	Chemical complementation of ER active ligands against RXRα	125
Figure 5.17	LBP of RARγ coupled with 9-cis retinoic acid	126
Figure 5.18	Chemical complementation ER active ligands against RARγ	127
Figure 6.1	Commonly used fluorescent tags	136
Figure 6.2	Classical Estradiol Modified Probes	137
Figure 6.3	Confocal images of GFP-chromophore inspired ERα ligands	139
Figure 6.4	AMI ligands bind selectively within *B. manjavacas* neonates	143

LIST OF SYMBOLS AND ABBREVIATIONS

3-AT	3-amino-1,2,4-triazole
ADE	adenine
AF-1	activation transcription factor-1(ligand independent)
AF-2	activation transcription factor-2 (ligand dependent)
ApoE	apolipoprotein E
AMI	Arylmethyleneimidazolone
ATP	Amino triphosphate
Arg	Arginine
AR	Androgen Receptor
BFP	Blue Fluorescent Protein
Ca^{2+}	Calcium ion
CAR	Constitutive Androgen Receptor
CFP	Cyan Fluorescent Protein
CMV	Cytomegalovirus
DAX1	Dosage sensitive sex-reversal AHC on chromosome X gene 1
DBD	DNA binding domain
DNA	Deoxyribonucleic Acid
DPN	Diarylpropionitrile
E2	Estradiol
ELISA	enzyme-linked immunosorbent assay
ER	Estrogen Receptor
ERα	Estrogen Receptor α
ERβ	Estrogen Receptor β

ERR	Estrogen Related Receptor
ESI	Electron spray ionization
eGFP	Enhanced Green Fluorescent Protein
FDA	Food and Drug Administration
FP	Fluorescent Protein
FRET	Fluorescence Resonance Energy Transfer
FRFP	Far Red Fluorescent Protein
FXR	Farnesoid X Receptor
GAD	Gal 4 Activation Domain
GBD	Gal 4 DNA binding domain
Gln	Glutamine
Gly	Glycine
GFP	Green Fluorescent Protein
GR	Glucocorticoid receptor
H12	Helix 12
HCl	Hydrochloric Acid
HDAC	Histone Deacetylase
HEK 293T	Human embryonic kidney cell 293T
His	Histidine
HLW	Histidine, Leucine, Tryptophan (media lacking these amino acids)
HRE	Hormone Response Element
HSA	Human Serum Albumin
HSP	Heat Shock Protein
LBD	Ligand Binding Domain
LBP	Ligand Binding Pocket

LXR	Liver X Receptor
MR	mineralocorticoid receptor
M	Molar
Met	Methionine
NCoR	Nuclear Receptor Corepressor
NL	No ligand
NLS	Nuclear Localization Signal
NMR	Nuclear Magnetic Resonance
OD	Optical Density
OFP	Orange Fluorescent Protein
PCG-1α	Principal Co-factor 1
PDB	Protein data bank
PG	Prostaglandin
PPAR	Peroxisome proliferators activated
PPT	propylpyrazole
PR	Progesterone Receptor
RAR	Retinoic Acid Receptor
RFP	Red Fluorescent Protein
RLU	Relative Light Unit
RNA	Ribonucleic Acid
RXR	Retinoid Acid Receptor
SAR	Structure Activity Relationship
Ser	Serine
SERBA	Selective Estrogen Receptor Beta Agonist
SERM	Selective Estrogen Receptor Modulator

SHP	Small heterodimer partner
SMRT	Silencing Mediator for Retinoid and Thyroid receptor
SNuRM	Selective Nuclear Receptor Modulator
SRC-1	Steroid Receptor Co-Activator 1
TMS	Tetra Methyl Silane
TFRET	Time Resolved Fluorescence Resonance Energy Transfer
TR	Thyroid Receptor
Tyr	Tyrosine
VDR	Vitamin D Receptor
YFP	Yellow Fluorescent Protein

SUMMARY

Nuclear receptors are ligand activated transcription factors that are widely distributed throughout the mammalians. There are 48 known human nuclear receptors within the body located in various systems. While some nuclear receptors can be located wholly within certain regions and tissues the clear majority are widely distributed, overlapping expression in the same locations. The role of nuclear receptors as transcription factors has caused them to be implicated in a vast number of diseases including metabolic, cardiovascular and neurological. The role of nuclear receptors in diseases and the potential to promote ligand activated transcription makes nuclear receptors of pharmaceutical significance. Currently it is estimated that 33% of nuclear receptors are targeted by the pharmaceutical industry resulting in ~20% of pharmaceutical development worldwide. The potential to control physiological responses via introduction of a ligand to nuclear receptors has continued the interest in development of new ligands to further the understanding of nuclear receptor behavior. The challenge nuclear receptors present is to develop ligands that are selective in targeting within families and among different classes of nuclear receptors.

At the core, ligand activated nuclear receptor modulation is chiefly centered around the relationship between the ligand binding pocket of the receptor and the ligand. Composed primarily of non-polar amino acid residues the ligand binding pocket is the cavity by which small hydrophobic molecules bind. Demonstrating large variance across classes of nuclear receptor and little divergence within families the ligand binding pocket serves as the focal point for targeting selectivity. Successful binding and thus receptor response is contingent upon the

ligand meeting criteria established by the ligand binding pocket such as satisfactory size/volume of the ligand and key ligand-receptor amino acid residue interaction. Research conducted by Katzenellenbogen was paramount in understanding the relationship between the estrogen receptors and its ligands. His established pharmacophore unraveled features for potential ligands that are exchangeable from those that are indispensable. The commercial success of estrogen receptor ligands has fueled the interest in not only understanding ligand-receptor binding interactions but its subcellular movement.

The Green Fluorescent Protein completely revolutionized the way in which cellular probing is conducted. The chromophore internally synthesized by the protein through a series of folding of amino acid residues afforded the opportunity to monitor cellular movements with the aid of fluorescence. Commonly utilized in visualization as a fusion protein, the GFP chromophore provided the ideal tool for understanding protein cellular movement and interaction. Simply put due to the chromophore that resides at the center, GFP provides the perfect technique for cellular probing.

Here in we report the use of GFP-chromophore inspired ligands for utility as estrogen receptor agonist. By utilizing the GFP chromophore skeleton as a template for ligand development the potential arises for the molecule to co-function as a receptor binder and probe. The use of the GFP-chromophore skeleton boasts several advantages in addition to synthetic amenable features the arymethyleneimidazolone core maintains the same frame work as Katzenellenbogen's proposed pharmacophore. Through the lens of Katzenellenbogen's consideration and the use of a simple but elegant synthesis a small library of GFP chromophore inspired arylmethyleneimidazolone ligands were synthesized, screened for selective estrogen receptor activation and tested both in vivo and in vitro for cellular probing applications.

Through this work we identified a set of 10 ligands that serve as agonist for the estrogen receptor. Although of the 10 ligands several demonstrate activation for both ERα and ERβ, a high degree of preference for ERα is observed. Of the ligands screened all estrogen receptor active ligands were nuclear receptor selective failing to activate other receptors such as RAR and RXR. Biological screening also uncovered a super agonist in CW32 that demonstrated the highest level of activation. Though a structure activity relationship model was established for top activators and additional generations synthesized no compound was found to be more active than CW32. While the majority of ligands displayed a preferential affinity for ERα ligand CW72 demonstrated complete specificity for ERα. All ligands were confirmed through TRFRET as binding in the same ligand binding pocket as estradiol. Computational modeling supports the rationale that the following three criteria governed the ligands ability to successfully bind: 1) hydrogen bonding network, 2) ligand size/volume and 3) molecular topology.

Embracing the ligands skeleton originating from the Green fluorescent chromophore ligands that demonstrated ER activation were visualized under confocal both with in vivo and in vitro systems. Several ligands successfully demonstrated the ability to turn on fluorescence in responds to binding in vitro. While other ligands failed to display fluorescence in conjunction with binding. Despite all binders displaying fluorescence this represent a class of ligand that can serve as a activator and probe.

CHAPTER 1

OVERVIEW

1.1 Motivation

Ligand design and target selection rest at the center of drug discovery and development however each stage presents its own set of challenges and obstacles [1]. The drug discovery process can be unsuccessful for multiple reasons including screening, potency and selectivity among the most common complications. These hardships are typically encountered early in the process such as during screening if available. To adequately evaluate the ligands ability to influence a biological system a model for assessing must be developed, however not all models include the target [2]. Exclusion of the target from the screening process at times can give an inadequate representation of results such as potency and selectivity [2]. Promiscuity in physiological response is often a result of the ligand interacting with alternative bio-factors in addition to the target. Ligand selectivity is of prime importance ideally the ligand should only activate the desired target. In the case of protein targets not only must selectivity be generated from similar classes of protein but diverging proteins within the class must be accounted for as well. The ligands ability to interact with a biochemical target is a critical early step in the process of promoting physiological response, more insight in the cellular behavior could yield creation of ligands with high degree of selectivity.

A better understanding of the ligand-target complex is often achieved with fluorescently tagged ligands for cellular tracking and visualization. This common practice provides insight into both ligand movement and target location within tissue [3]. Once the ligand complexes with its target

additional information can be gained about the ligand-target complex and its post translational movement. The insight provided by fluorescently labeling the ligand can aid in re-designing of ligands that are more target specific and possibly tissue selective potentially resulting in diminishing of undesired biological responses. While tracking and visualization through the addition of a fluorescent tag is resourceful and insightful it is not without its drawbacks. The addition of a tag increases the ligands natural size which may adversely affect solubility of the ligand, ability to interact with binding sites of the target, and biological uptake of the ligand. Biological targets act in concert with larger systems, many undergo folding or some form of change after coupling with a ligand such proteins. The additional size added from the tag may restrict the protein or target from undergoing its natural conformation change thus altering its physiological response or biological process.

The intent of this thesis is to convey the use of arylmethyleneimidazolone skeleton inspired by the green fluorescent protein chromophore as ligands that are potent, selective and provide tracking without a tag. Previously conducted work by Dr. Anna Duraj-Thatte [4] and Dr. Anthony Baldrige [5] identified nuclear receptors as a class of targets, for the scope of this thesis the estrogen receptor was selected as target of interest. By utilizing the chromophore of the green fluorescent protein as the backbone structure for our ligand design we adopt a structure with known fluorescence capability and strong precedence for uptake by proteins since chromophore formation occurs within the protein [3,6-8]. The overall driving question was can the chromophore inspired skeleton be an activating ligand, be selective, and maintain intrinsic photo-physic properties that could be used for visualization? To test this hypothesis 3 research aims were devised.

1.1.1 Designing of a potent estrogen receptor agonist

Research Aim 1: Design of a potent estrogen receptor agonist

Initial efforts in identifying a lead for this class of ligand started with results of an unpublished screening by our lab [4]. A library of GFP chromophore inspired ligands was screened for estrogen receptor activity via two methods yeasts three hybrid assay and luciferase assay. While the results indicated activity in seven ligands only two demonstrated activation significant enough to prompt further investigation as a lead compound. Of the over 100 compounds screened for estrogen receptor α AB 18 shown in figure 1.1 was selected as the lead due to an activation level 67% that of estradiol, the indigenous ligand.

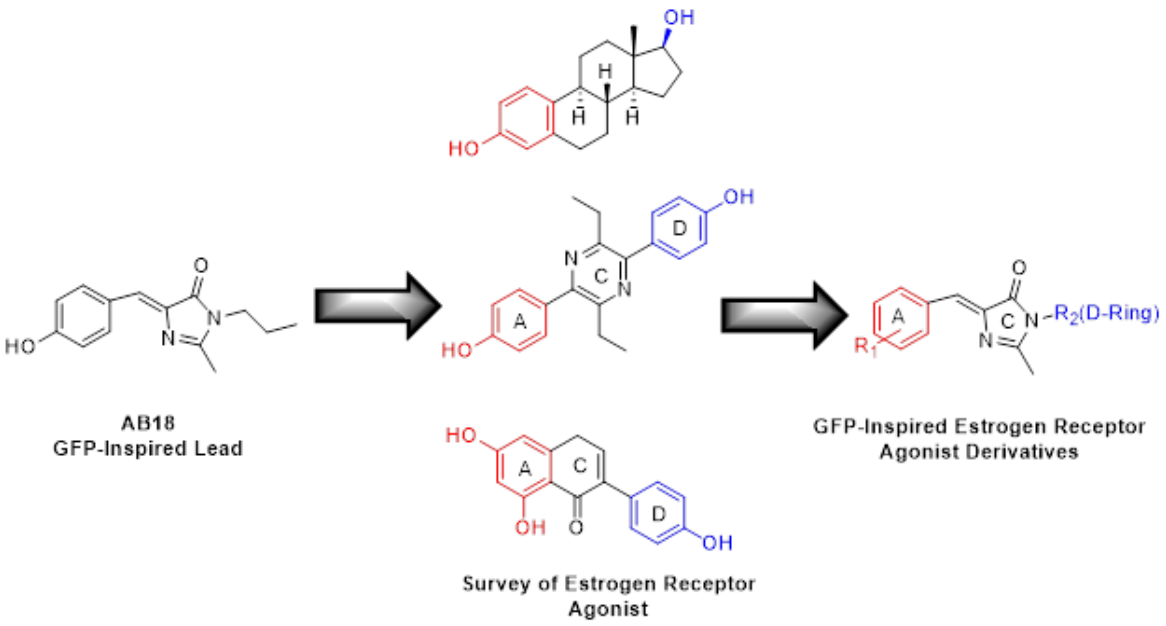


Figure 1.1 Lead Development for Research Aim 1. Development of Aim 1through surveying of known estrogen receptor agonist to build commonalities highlighted in red and blue into the GFP-inspired estrogen receptor agonist.

Using AB18 as a lead, structural characteristics of known estrogen receptor ligands were surveyed to understand pre-requisites for binding. The objective was to design and synthesize

ligands that would not only bind but maximize receptor-ligand interactions in a hope to enhance potency. This stage is set to be an iterative process between ligand generation and ligand screening with the goal of developing a functional structure activity relationship ultimately resulting in a highly potent estrogen receptor ligand.

1.1.2 Promoting selectivity within isotypes and among nuclear receptors

Research Aim 2: Promote selectivity within estrogen receptor isotype and among other nuclear receptors.

To test the hypothesis that this class of ERα ligand can be isotype and receptor selective all ligands deemed active during aim one will be screened for selectivity first within the estrogen receptor isotype and secondly among other nuclear receptors. The approach will be to compare the ligand binding domains of estrogen receptor subtype β with that of estrogen receptor α for features that could be exploited to promote selectivity. Structural differences of ligands will also be examined with attention devoted to ligands that have demonstrated isotype selectivity. The information learned in aim 1 in conjunction with revelation of estrogen receptor β ligand binding domain will be used to design and synthesize additional compounds aimed to promote selectivity while maintaining high activation.

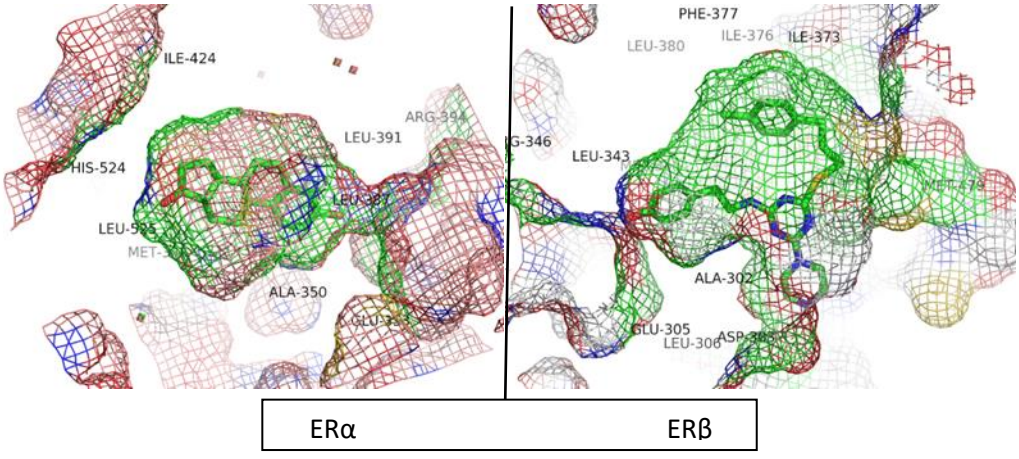


Figure 1.2 Estrogen receptors ligand binding domains complexed with ligand. Diversity can be seen between the estrogen receptors ligand binding domain and interactions with ligands providing an avenue for design of selective ligands. Labels for residues within the ligand binding pocket 4 Angstroms away from the ligand are included. The image on the left shows the ligand binding domain for ERα complexed with estradiol. The right illustrates ERβ ligand binding domain. Figure generated using Pymol with proteins from PDB

In addition to ERβ, selectivity will be pursued across various nuclear receptor families such as retinoic x receptor α and retinoic acid receptor γ chosen for heterodimerization ability [9] and role in cancer [10] respectively. The approach will employ exploiting the variance that exist between ligand binding domains of different nuclear receptor families. Only ligands showing activation through aim 1 or initial phase of aim 2 will be screened in chemical complementation for retinoic x receptor α and retinoic acid receptor γ. Moving forward only active ligands will be utilized in proceeding aims.

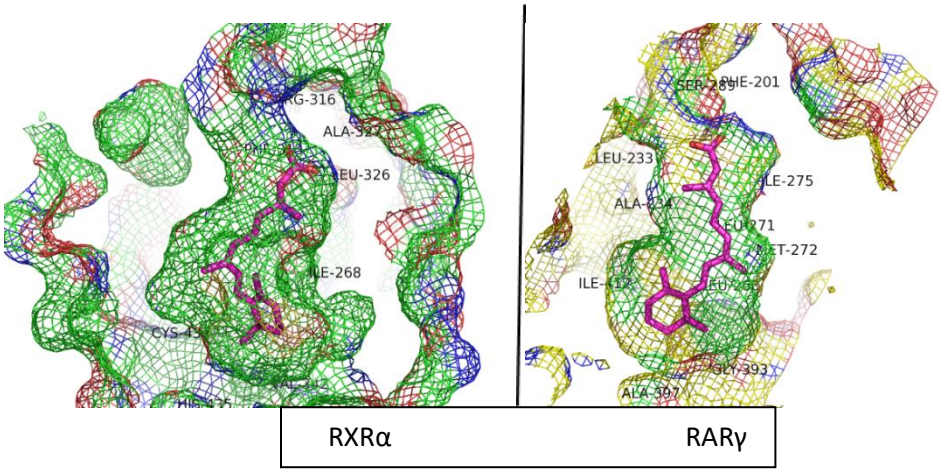


Figure 1.3 Ligand binding domains of additional nuclear receptors complexed with ligands. Image on the left shows retinoid X receptor ligand binding domain complexed with 9-cis retinoic acid while image on the right shows retinoic acid receptor complexed with 9-cis retinoic acid. By exploiting differences within the pockets of nuclear receptor ligand selectivity can be generated for additional categories of receptors. Figure generated using Pymol with proteins from PDB

1.1.3 Accessing the role of florescence and activation

Research Aim 3: Access the role of florescence and activation with chromophore inspired ligands.

To investigate the hypothesis that this class of ligand can be utilized as cellular probes, ligands identified as active will be subject to visualization via confocal microscopy. Visualizing active ligands both in the absence and presence of the estrogen receptor α ligand binding domain and cells will provide the opportunity to examine if fluorescence can be observed in active ligands only upon binding or is fluorescence achieved independent of a binding environment. The successful accomplishment of this aim will promote a ligand that not only binds but can be used for cellular tracking of the estrogen receptor α without traditional fluorescent labels used for estrogen receptor ligand agonist.

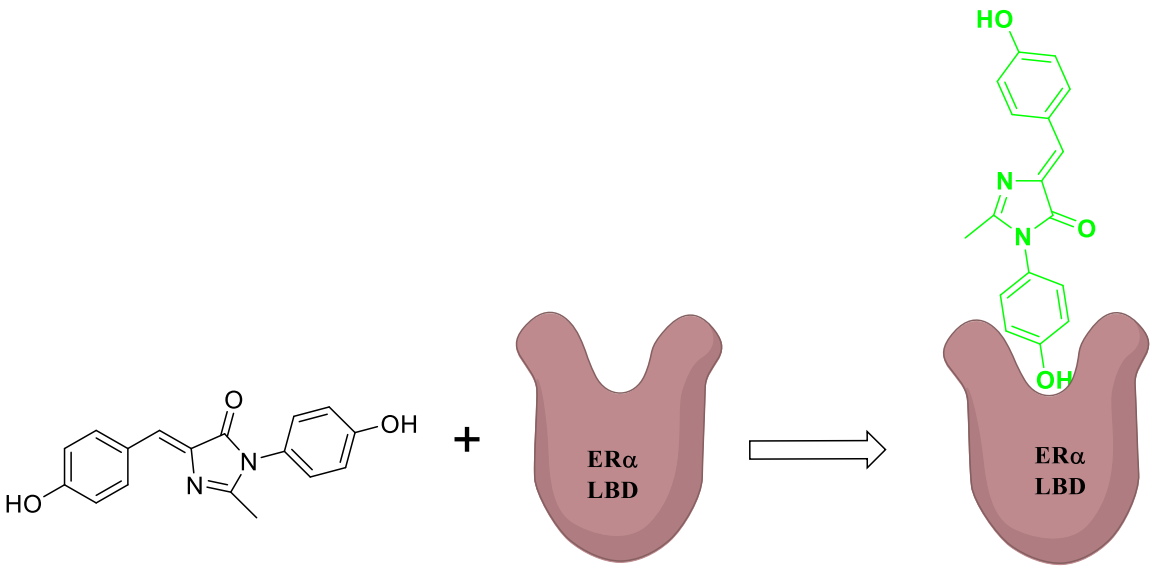


Figure 1.4 Turn on of GFP inspired chromophore fluorescence via binding to estrogen receptor α. In the absence of the protein ideally the chromophore should not be activated. Upon binding of the ligand torsional strain should re-activate intrinsic fluorescence properties tying binding and fluorescence together.

The overall project presents several points of novelty while other similar ring systems have been utilized including 2-Arylindene-1-ones [11] and simple benzamides [12] to our knowledge this is the first class of green fluorescent protein inspired ligands for the estrogen receptor α agonist. Likewise, while green fluorescent protein has significant applications as a molecular probe the

ideal of utilizing the chromophore in the absence of the protein as a cellular probe and as nuclear receptor activator strengthens the originality of this research.

1.2 References

1. Wallace, Owen B.; Richardson, Timothy I.; Dodge, Jeffrey A. *Current Topics in Medicinal Chemistry*, **2003**, 3, 1663-1680.

2. Brown, D. *Drug Discovery Today*. **2007**, 12, 1007-1012

3. Tsien, R.Y. *Annu. Rev. Biochem.* **1998**,*67*, 509-544

4. Duraj-Thatte, A. FLUORESCENT GFP CHROMOPHORES AS POTENTIAL LIGANDS FOR VARIOUS NUCLEAR RECEPTORS **2012**. Georgia Institute of Technology

5. Baldridge, A. SYNTHESIS, PHOTOPHYSICS, AND APPLICATON OF FLUORESCENT PROTEIN CHROMOPHORE ANALOGS. **2011**, Georgia Institute of Technology.

6. Barondeau, D.P; Putnam, C.D.; Kassmann, C.J.; Tainer, J.A; Getzoff, E.D. *Proc. Natl. Acad. Sci. U. S. A*., **2003**, *100*, 12111–12116.

7. Cubitt, A. B.; Heim, R.; Adams, S. R.; Boyd, A. E.; Gross, L. A.; Tsien, R. Y. *TIBS*, **1995**, *20*, 448-455.

8. Heim, R.; Prasher, D. C.; Tsien, R. Y. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 12501-12504

9. Hinrich Gronemeyer, Jan-Åke Gustafsson, Vincent Laudet, *Nature Reviews.***2004**, 3.

10. S. Chandraratna. *J. Med. Chem.* **1995**, 38, 4764-4767.

11. Filippo Minutolo, Marco Macchia, Benita S. Katzenellenbogen, and John A. Katzenellenbogen, *Medicinal Research Reviews*, **2011**, 3, 364-442.

12. Caldarelli, Antonio; Minazzi, Paolo; Canonico, Pier Luigi; Genazzani, Armando A.; Giovenzana, Giovanni B. *Journal Enzyme Inhibition and Medicinal Chemistry*. **2013**,1, 148-152.

CHAPTER 2

GREEN FLUORESCENT PROTEIN

2.1 Historical Perspective

Since its discovery and identification, the green fluorescent protein has catapulted into the spotlight and become one of the most widely explored and utilized proteins in cell biology and biochemistry [1]. Pioneered by several key figures the revolution induced by GFP in fluorescent proteins as biological tools resulted in the awarding of the Nobel Prize in chemistry in 2008 to select foundational contributors.

Discovered in 1961 by Osamu Shimomura [2] from the jelly fish *Aequorea aequorea* (*Aequorea victoria*) GFP was initially described as a green fluorescence that occurred when Aequorea where irradiated with ultraviolet light [3]. Soon after its initial discovery GFP was identified as a protein [4] and renamed from "green protein" to its current nomenclature green fluorescent protein [5]. In 1979 the structure of the chromophore that gave rise to the fluorescence was characterized [6].

Aequorea Victoria photo courtesy of Glebstock-Fotolia

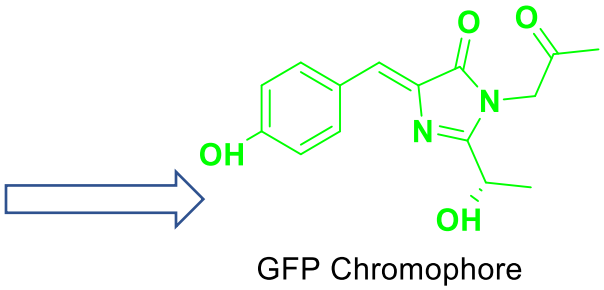


Figure 2.1 Aequorea victoria displaying fluorescence. GFP is responsible for the observed fluorescence in jellyfish. At the center of GFP's ability to fluoresce reside the above chromophore.

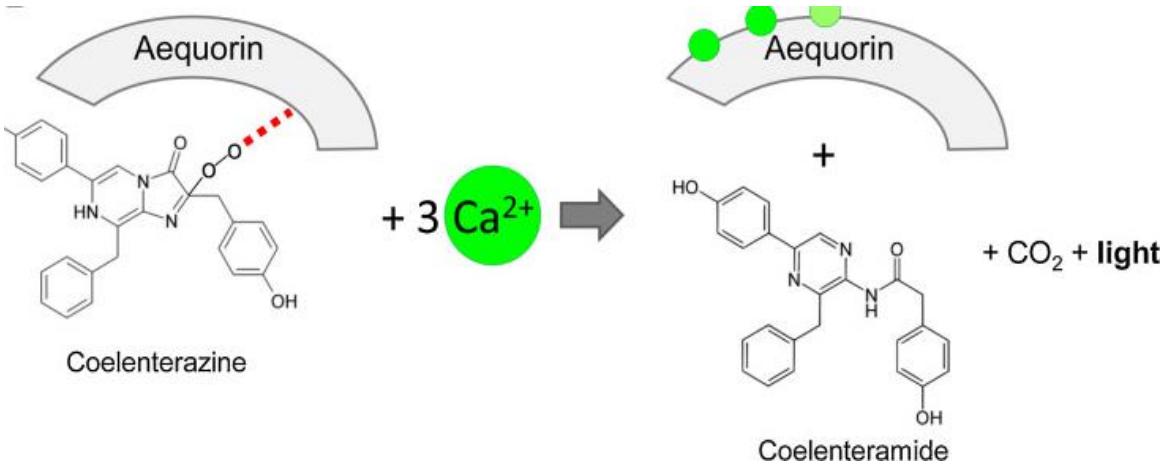
Accumulation of small amounts of GFP afforded Shimomura the opportunity to conduct a series of experiments including enzymic digestion, peptide isolation and purification. Spectroscopic analysis in conjunction with comparison of absorption spectra of previous compounds Shimomura synthesized, ultimately led to the deduction of the chromophore structure. Characterization of the chromophore revealed that fluorescence was not a result of a complex or association of a secondary organism, but amino acids encoded in the primary sequence of the protein that allowed the protein to autocatalyze the formation of its own fluorophore. This discovery cleared the path for the possibility of cloning the protein without infringing upon its fluorescent properties. In 1985 aequorin was cloned and expressed shortly thereafter GFP was cloned [7, 8, 9, 10] which expedited the progression of GFP. Over the next several years' advancements in GFP discoveries flourished including the confirmation of Shimomura proposed chromophore structure and flanking amino acids [11], GFP expression in E. coli [12], and crystal structure of wild type GFP [13]. In the years to follow the discovery of GFP and its chromophore paved the way for discoveries of additional fluorescent proteins [14] which will be discussed later in this chapter. Cloning and expression eased the accessibility of GFP thus encouraging novel applications in cell biology and visual development with pioneering studies being conducted in GFP lasers [15] and cellular tracking.

2.2 Green Fluorescent Protein

Bioluminescence can be defined as the emission of light by a living organism as a resultant of a chemical reaction [16]. The production of bioluminescence can broadly be described as occurring *via* the oxidation of a substrate termed the luciferin by an enzyme referred to as the luciferase. While bioluminescence is not confined to any specific environment it is estimated that

90% of deep sea marine life produce bioluminescence [16]. Observed in bacteria, algae, crustaceans, jellyfish, sea stars and squids it is postulated that the light producing reaction serves as a survival tool to lure prey, elude predators, or allure a mate in all bioluminescence can certainly be viewed as elegant communication.

Bioluminescence has been studied in many marine organisms among the most notable is the jellyfish *Aequorea victoria.* Reported in 1955, *Aequorea victoria* fluoresced green when irritated with ultraviolet light, later it was discovered that luminescence and green fluorescence could be attributed to two proteins Aequorin and green fluorescent protein [3]. Aequorin is a calcium sensitive protein probe characterized by two features its calcium binding region and coelenterazine its prosthetic group [17]. Upon binding of calcium ions conformational changes occur within the apoaequorin (luciferase) causing the oxidation of the coelenterazine substrate (luciferin) [17]. The transition of the substrate from its excited state (coelenteramide) back to its relaxed state produces emission of a blue light [17]. In the absence of the green fluorescent protein the blue light is observed however in the presence of GFP green fluorescence is observed such as in *Aequorea victoria* [18]. No evidence of a binding complex between aequorin and GFP have been observed however a radiation less energy transfer occurs from the apoaequorin-calcium-coelenterazine complex to GFP resulting in a blueshift and subsequently green emission [19].



Francesco Moccia, Alessandro Bertoli, M. Catia Sorgato.

Figure 2.2 Aequorin- Coelenterazine Mechanism. The figure above illustrates aequorin binding to coelenterazine and calcium generating coelenteramide and light. Blue light in the absence of GFP or green light in the presence of GFP.

2.2.1 Green Fluorescent Protein Discovery

The reporting of luminescence in jelly fish dates back to 50 A.D. by Pliny the Elder [20] where glowing slime of *Pulmo marinus* was described. Support for this observation followed with the reporting of yellow-green fluorescent masses located in the marginal canal of *Aequorea* [3]. Fluorescence was described as lasting 0.3-1.5 seconds with intensity reaching a maximum in less than a second. In addition to response to stimuli the study examined location of *Aequorea's* fluorescence sighting that luminescence is an intracellular process. Critical thinking and experimentation, six years after the initial reporting of the fluorescence resulted in extraction and purification of the protein responsible for *Aequorea* fluorescence [2]. In addition to isolation of the primary photoprotein a secondary protein was isolated in trace amounts. This secondary protein exhibited green fluorescence and was initially termed green protein before receiving the permeant name of green fluorescent protein seven years after its isolation [7].

2.2.2 Green Fluorescent Protein Structure

Composed of 238 amino acids (27KDa) [21] the green fluorescent protein is made up of 11 β strands aligned as a β-barrel with a single α-helix running diagonally through the center. Physically the β-barrel has a diameter of ~30A and a length of about 40A with small sections of α-helices forming caps on the ends of the barrel [22]. The nearly symmetrical cylinder serves as a protective unit for the fluorophore often giving rise to the moniker "light in a can". Investigations have indicated that GFP resistance to heat, pH, proteases, and denaturants can be attributed to the compaction of the β-barrel [23]. The architectural skeleton of GFP shows a compact domain where the strands comprising the β-sheet are tightly fitted, in conjunction with the α-helix lids and loops GFP's casing gives rise to several significant characteristics and an intriguing micro-environment.

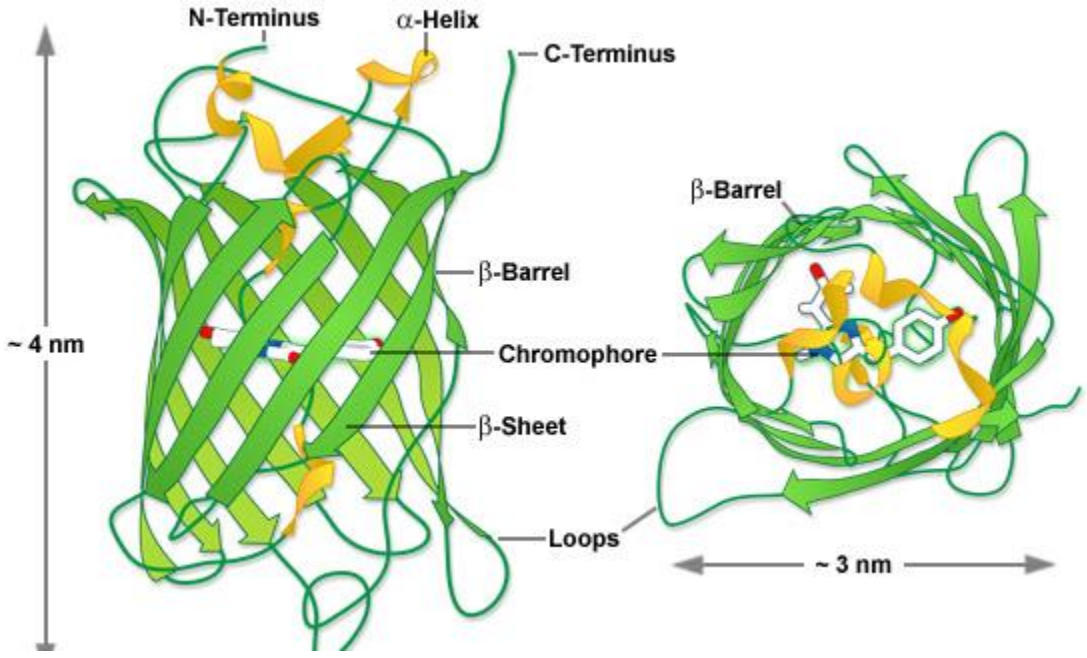


Photo courtesy of Carl Zeiss Microscopy

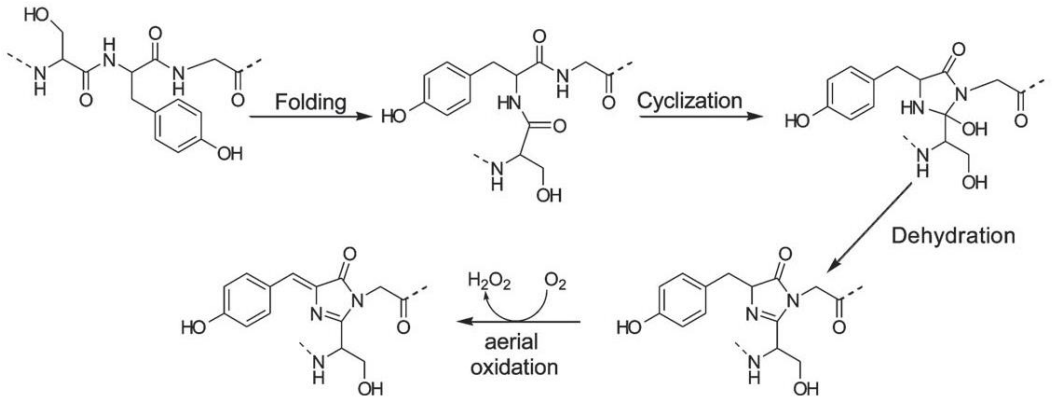
Figure 2.3 Structure of GFP. The above illustration shows the β-barrel of GFP with the chromophore located in the center.

2.3 Synthesis of GFP Chromophore

The paramount feature of GFP, its chromophore resides near the geometric center of the compact cylindrical protein along the α-helix. The fluorophore is derived from peptides of the proteins own backbone structure. Generated from a triplet of amino acids Ser-65, Tyr-66, Gly-67 the driving force for chromophore formation is the stability acquired from the folding of the protein [1]. Once the protein advances to its near mature conformation the chromophore initiation process is set in motion by an approximately 80^{o} bend within the tripeptide sequence [24]. The consequence of this bend disposes of a series of hydrogen bonds involving several main chains, now allowing for the repositioning of main chain residues near reactive sites. The widely accepted mechanism for the occurrence of this autocatalytic post translational multistep process was proposed by Tsien [25,26]. During the first step of the conjugation reaction, the amide nitrogen of Gly-67 attacks the carbonyl carbon of Ser-65 to promote internal cyclization. Cyclization is followed by dehydration to form an imidazolin-5-one intermediate [25]. Finally, during the rate determining step, molecular oxygen catalyzes dehydrogenation along the Cα–Cβ bond of Tyr-66 to form the fully conjugated ring structure yielding the 4-(p-hydroxybenzylidene)-imidazolidin-5-one or p-HBI chromophore. Though the multi-step process depends on molecular oxygen the protein backbone cyclization can occur under anaerobic conditions indicating that the proteins local environment assist in p-HBI chromophore formation.

The environment around the chromophore contains both apolar and polar amino acid side chains in addition to water molecules forming a hydrogen bonding network [19]. The residues within the surrounding pocket provide a scaffold for functional group catalysis specifically Arg-96 and Gln-222. Wachter proposed that both Arg-96 and Glu-222 residues are directly implicated in the cyclization and oxidation reactions [27]. Arginine-96 is believed to play a role through its ability

to hydrogen bond with carbonyl oxygen of Ser-65 thereby increasing the electrophilicity of the carbonyl carbon and promoting nucleophilic attack from Gly-67 [28]. Glutamic acid 222 is postulated to function as a base via carboxylate, this proton abstraction occurs from the α-carbon of Tyr-66 during the oxidation process [29]. In addition, Barondeau proposed that Arg-96 together with Tyr-66 promotes α-enolate intermediate formation prior to oxidation reaction [30]. Various mutagenesis suggests that the Gly-67 is required for fluorophore formation [21, 31] however central tyrosine can be replaced by any aromatic containing residue and still yield fluorescent protein [32]. The protein's local environment not only plays an assistive role in chromophore formation it also plays a pivotal role in the chromophores orientation. Existing in a cis-configuration causes the conjugated rings to assume a nearly co-planar orientation and the p-HBI chromophore is held in rigid form by an extensive hydrogen bonding network. This hydrogen bonding network composed of neighboring amino acid residues prevents twisting or tilting of the chromophore while in the first electronic excited state [21]. Studies have suggested that in the absence of constraint by the proteins environment rapid dispersal of the excited state energy would occur as thermal or rotational energy resulting in a non-radiative decay to the ground state [33].



Scheme 2.4 Autocatalytic Formation of GFP-Chromophore. The above scheme illustrates the biosynthetic pathway of the GFP chromophore

2.4 GFP Photochemical Properties

The innate relationship between the protein architecture and its chromophore formation has several translations exhibited in the proteins photochemical behavior. Wild type GFP exhibits two absorptions, a major peak at 398 nm and a minor shoulder at 475 nm with the different peaks being attributed to changes in protonation of the chromophore [33]. GFP's chromophore can exists in two ionization states a neutral and anionic form. Although separate ionization states are responsible for two absorption frequencies, both exhibit approximately the same emission spectra frequency. Excitation at the major absorption frequency (neutral form) results in an emission at 508 nm while excitation at the latter (anionic form) yields a maximum of 503 nm [33]. GFP single emission is the resultant of a common electronic exited state of the fluorophore. Upon excitation the neutral ionization state goes to near zero becoming a strong acid this is only achievable by proton transfer through a pre-organized hydrogen bonding network. This excited state proton transfer (ESPT), which has been investigated in studies rapidly generates the excited state anion which emits green light [34-38].

Mutagenesis studies have been employed to explore the effect of the local protein environment on the photochemical properties of GFP. Studies have demonstrated that mutations within the regions adjacent to the fluorophore can potentially lead to significant wavelength shifts however the majority result in loss of fluorescence [22]. The critical nature of Gly-67 for chromophore formation safe-guards this residue from mutation nonetheless mutations in tyrosine and serine have been investigated. Mutation of central tyrosine to phenylalanine or histidine results in a band shift but predominately yields a significant loss in intensity [31]. Intriguingly the mutation of serine to threonine results in an increase in fluorescence intensity [26,40]. The proposed explanation for the increase in intensity is a reduction in collisional quenching in the excited

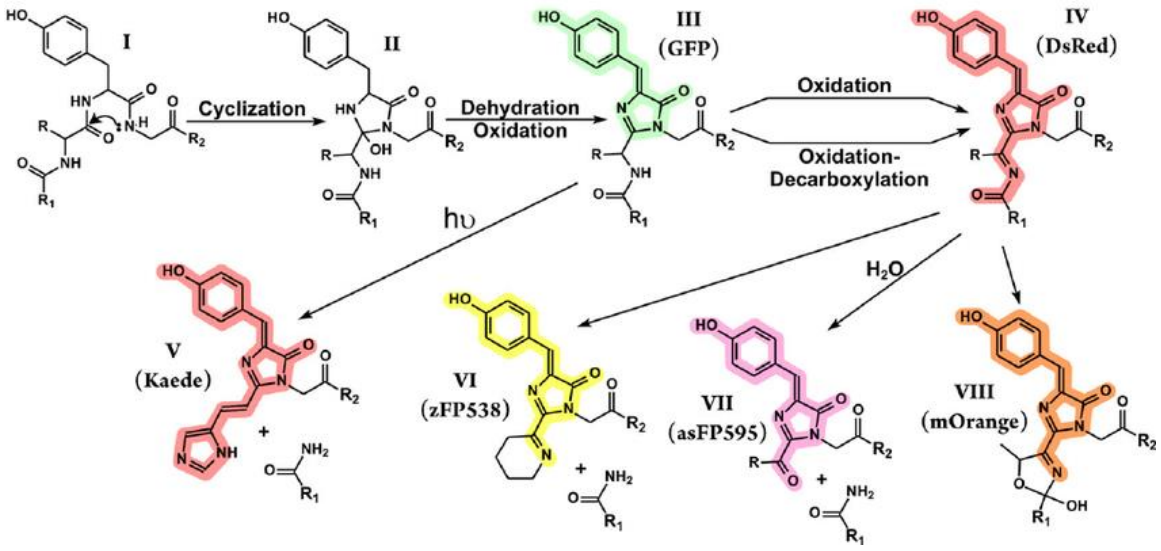
state. This reduction is derived from the interactions of better packing within the interior of the protein because of the inclusion of the methyl moiety from threonine substitution [22]. Increased packing provides less exposure to the fluorophore providing a possible explanation as to reports that mutated serine to threonine or enhanced GFP (eGFP) is more resistant to photo bleaching than the wild type species [25].

Overall the photo properties of GFP can be modulated by its local environment given rise to GFP's increasing utility as a cellular probe.

2.5 Fluorescent Proteins

Insight into the mechanism of the GFP chromophore formation and the role of surrounding amino acid residues assisted in the uncovering of other fluorescent protein (FP) GFP variants. Understood to be divided into seven spectral divisions based upon the emission maxima GFP variants are categorized as the following: blue fluorescent protein (BFP), cyan(CFP), yellow (YFP), orange (OFP), red (RFP) and far-red (FRFP). The p-HBI chromophore or variation of exist in all naturally occurring FP's, however several factors determine fluorescent characteristic. Diversity in spectral profile is achieved by direct chemical modification of the p-HBI, via mutations to the residues involved in chromophore formation or residues in the protein microenvironment. Mutations in the central Tyr-66 in GFP gives birth to two distinctive variants BFP and CFP. Introduction of a histidine residue at Tyr-66 results in the formation of an imidazole ring during chromophore formation as opposed to a phenol moiety which correspondingly emits in the blue range. Substitution of identical Tyr-66 to tryptophan leads to the formation of an indole in the chromophore resulting in cyan fluorescence. Additional GFP variant subfamilies such as orange and red fluorescent proteins utilize the p-HBI chromophore as an intermediate where subsequent catalysis forms linkage with the polypeptide back bone and

extends conjugation. Mutations in the protein local environment also give rise to GFP variants through associative interactions. The mutation of threonine 203 to tyrosine results in π-π stacking interactions between the residue and the phenolate ion in the chromophore. This interaction thus stabilizes the dipole moment of the chromophore resulting in yellow emissions which consequently turns out to be the brightest of the FP families [40]. The mutations and affiliated fluorescent protein emissions are displayed in figure 2.5, the available variety of colored proteins increases the utility within cellular applications.



Pakhomov,A. Chemistry & Biology,2008,15,755-764.

Figure 2.5. Color Variants of Fluorescent Proteins. The above scheme illustrates the diversity in colors that arise from modifications to GFP chromophore.

2.6 References

1. Tsien, R.Y. *Annu. Rev. Biochem.* **1998**,*67*, 509-544

2. Shimomura, O. *Angew. Chem. Int. Ed.* **2009**, *48*,5590-5602.

3. Davenport, D. *JAC Nicol:. Proc. R. Soc. London, Ser. B.***1955***,144*,399-411.

4. Shimomura, O; Johnson, FH; Saiga, Y. *J. Cell. Comp. Physiol.* **1962***,59*, 223-29.

5. Hasting, JW; Morin, JG. *Biol. Bull*, **1969**, *137*, 402.

6. Shimomura, O. *FBS Letters*, **1979**, *104*, 220-222.

7. Inouye, S; Noguchi, M; Sakaki, Y; Takagi, Y. *Proc. Natl. Acad. Sci USA*, **1985**, *82*, 3154-3158.

8. Inouye, S; Sakaki, Y; Goto, T; Tsuji, F.I. *Biochemistry*, **1986**, *25*, 8425-8429.

9. Prasher, D; McCann, R.O.; Cormier, M.J. *Biochem. Biophys Res Commun,* **1985**, *126*, 1259-68.

10. Prasher, D.C.; Eckenrode, V.K.; Ward, W.W.; Prendergast, F.G.; Cormier, M.J. *Gene,* **1992**, *111*, 229-233.

11. Cody, CW; Prasher, D.C.; Westler, WM; Pendergast, F.G.; Ward, W.W. *Biochemistry*, **1993**, *32*, 1212-18.

12. Chalfie, M; Tu, Y; Euskirchen, G; Ward, W.W.; Prasher, DC. *Science* **1994**, *263*, 802-805.

13. Mormo, AB; Cubitt, K; Kallio, LA; Tsien, RY; Remington, SJ. *Science*, **1996,** *273*,1392-95.

14. Matz, MV; Fradkov, AF; Labas, YA; Savitisky, AP; Zaraisky, AG; Markelov, ML; Lukyanov, SA. *Nature Biotech.* **1999***, 17*, 969-73.

15. Gather, MC; Yun, Sh. *Nature Photonics*, **2011**

16. www.sciencedaily.com/terms/bioluminescence.htm

17. Shimomura, O; Musicki, B; Kishi, Y. *Biochem. J*, **1988**, 251, 405-410.

18. Kendall, J.M; Badminton, M.N. *TIBTECH*, **1998**, 16, 216-224.

19. Zimmer, M. *Chemical Reviews*, **2002**, *102*, 3, 759-781.

20. www.scripps.ucsd.edu

21. Vedangi, S; Newman, R; Zhang, J. *Chem. Soc. Rev.,* **2009**, *38*, 2852-2864.

22. Yang, F; Moss, L.G.; Phillips, G. *Nature Biotechnology*, **1996**, *14,* 1246-1251.

23. Klaithem, M; Alkaabi, A.Y.; Salman, A. *Applied Biochemistry and Biotechnology*, **2005**, *126*.

24. Barondeau, D.P; Putnam, C.D.; Kassmann, C.J.; Tainer, J.A; Getzoff, E.D. *Proc. Natl. Acad. Sci. U. S. A*., **2003**, *100*, 12111–12116.

25. Cubitt, A. B.; Heim, R.; Adams, S. R.; Boyd, A. E.; Gross, L. A.; Tsien, R. Y. *TIBS*, **1995**, *20*, 448-455.

26. Heim, R.; Prasher, D. C.; Tsien, R. Y. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 12501-12504.

27. Wachter. *Acc. Chem. Res.* **2007**, *40*, 120-127.

28. Wood, T.I; Barondeau, D.P; Hitomi, C; Kassmann, C.J; Tainer, J.A; Getzoff, ED. *Biochemistry,* **2005**, *44*, 16211-16220.

29. Sniegowski, J.A.; Lappe, J.W.; Patel, H.N; Huffman, H.A.; Wachter, R.M. *J. Biol. Chem.* **2005**, *280*, 26248-26255

30. Barundeau, D.P; Tainer, J.A; and Getzoff, ED. *J. Am. Chem Soc.* **2006**, *128*, 3166-3168.

31. Delagrave, S.; Hawtin, R.; Slilva, C.; Yang, M.; Youvan, D. *Bio-Technology*, **1995**, *13*, 151-154.

32. Remington, S.J. *Current Opinion in Structural Biology*, **2006**, 16, 714–721.

33. Remington. **2014** The Flourescent Protein Revolution,Chapter 2., Lessons learned from structural studies of fluorescent protein.,

34. Chattoraj, M.; King, BA.; Bublitz, GU.; Boxer, SG. *Proc Natl Acad Sci USA*, **1996**, *93*, 8362-8367.

35. Breje, K.; Sixma, TK; Kitts, PA.; Kain, SR.; Tsien, RY.; Ormo, M.; Remington, SJ. *Proc Natl Acad Sci USA,* **1997**, *94*, 2306-2311.

36. Palm, GJ; Zdanov, A; Gaitanavis, GA; Stauber, R; Pavlakis, GN; Wlodawer. *Nat Struct Biol,* **1997**, *4*, 361-365

37. Stoner-MaD; Jaye, AA; Matousek, P; Towric, M; Meech, SR; Tonge, PJ. *J. Am Chem Soc*, **2005**, *127*, 2864-2865.

38. S. Meech, Chem. Soc. Rev., 2009

39. Heim, R; Cubitt, A; Tsien, R. *Nature*, **1995**, *373*, 663-664.

40.Shaner, N.C.; Patterson, G.H.; Davidson, M.W. *J. Cell Sci*., **2007**, *120*, 4247–4260.

CHAPTER 3

NUCLEAR RECEPTORS

3.1 Nuclear Receptors Overview

Essential in a series of physiological and biological process nuclear receptors are ligand regulated transcription factors. Superfamily of nuclear receptors serve as on/off switches with the ability to modulate target gene transcription. These “molecular switches” after ligand binding initiate specific genetic sequences that regulate extracellular environments from intracellular communication. Nuclear receptors reside in a wide variety of tissue and organ distribution within the body with diverse receptors being expressed at different levels in different organelles. At a subcellular level some nuclear receptors translocate, initially residing in the cytoplasm in an unbound state, upon ligand binding the receptor permeates the nucleus. Except for the estrogen receptors that exist in both bound and unbound state in the nucleus, nuclear receptors are cytosolic. The research that encompasses nuclear receptors has developed into a broad spectrum of topics ranging from structural & functional analysis, molecular mechanism and ligand modulation.

The first nuclear receptor was identified biochemically in the 1960’s (Jensen), its identification subsequently led to the first cloning of a nuclear receptor, glucocorticoid receptor (Evans) [1]. Initially isolated and identified via laborious laboratory techniques, the knowledge that nuclear receptors could be isolated without the ligand increased the rate of identification [2]. Confirmed in 2001 via the human genome sequence the 48 nuclear receptor genes located in mammalians are the complete nuclear receptor genome [3]. Among the 48 identified nuclear receptors only 24

have either endogenous ligands or identified ligands, those receptors lacking known ligands are termed orphan receptors [4]. More commonly the superfamily of nuclear receptors has been divided into subfamilies predicated upon the class of ligand they are regulated by. The bulk of the nuclear receptor superfamily is predominately regulated by small hydrophobic molecules from diverse classifications including steroids, bile acids, and hormones. Incorporated in the subfamilies are the steroid receptors, retinoid (RXR heterodimer) receptors and xenobiotic receptors. The subfamily of steroidal receptors composes the glucocorticoid receptor (GR), mineralocorticoid receptor (MR), progesterone receptor (PR), androgen receptor (AR), and the estrogen receptor (ER). This set of receptors of which many have isomeric forms are activated by steroidal ligands both synthetic and natural. The next subfamily commonly activated by fatty acids, oxysterols, bile acids, vitamins and hormones are the retinoidal or RXR heterodimer receptors. This category of receptor houses the peroxisome proliferators activated (PPAR), liver x receptor (LXR), farnesoid x receptor (FXR) and the retinoid x receptor (RXR). Though divided into categories and sub-divisions and housing cellular controls for various processes these modulators harbor inherent features.

3.2 Nuclear Receptor Structure (Regions and Domains)

A unifying feature of the nuclear receptor superfamily is that each receptor consists of an assembly of functional modules. Nuclear receptors are composed of several characteristic domains commonly denoted as A-F regions [5,6] as illustrated in Fig. 3.1. The A/B region also noted as the N-terminal domain houses an independent-ligand activation transcription factor (AF-1). This region displays a large degree of variance not only among nuclear receptors but also within isoforms. In fact, the A/B region is one of the least conserved among receptors,

however numerous studies indicate a commonality in the ability to regulate via phosphorylation [7-9]. To date no three-dimensional structure has been solved for this region. Residing within the C-region is the DNA binding domain (DBD). The DBD one of two focal points of the nuclear receptor is connected to the ligand binding domain (LBD) by a hinge designated as region D. The E region houses the activation factor 2 (AF-2) as well as the LBD, the hallmark of the nuclear receptor. Lastly the F region is composed of the C-terminal domain which is highly variable in sequence between various receptors. The merger of these functional modules creates a harmonic molecular machine capable of decoding ligand information to turn on transcription.

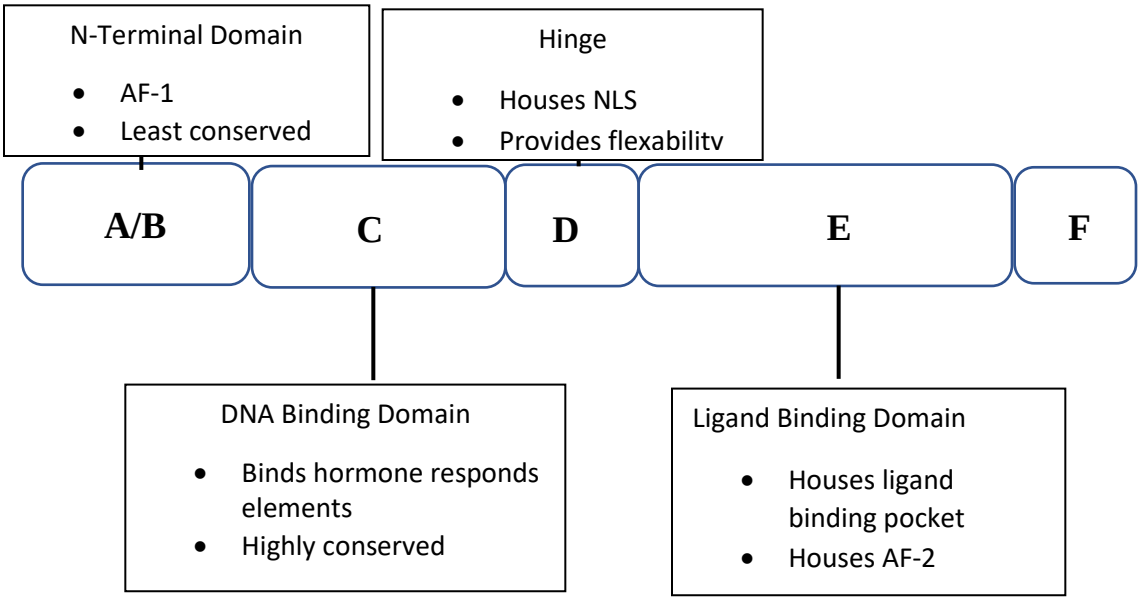


Figure 3.1 Nuclear receptors regions-domains. Structurally nuclear receptor shares 5 common domains illustrated above are the domains with defining characteristics.

3.2.1 Activation transcription factor (AF-1)

AF-1 varies in length and sequence between receptor groups and families. It is recognized by various accessory proteins termed promoters [10] however demonstrates specificity for cell, DBD and promoter interaction.

3.2.2 DNA binding domain

The DNA binding domain is one of two extensively studied focal points within the nuclear receptor. The predominate role of the DBD is to recognize and bind to specific DNA sequences termed hormone response elements (HRE). These short sequences of DNA exist within the promoter of the gene and allow for regulation of transcription [5]. The DNA binding domain is highly conserved among receptors being observed in all nuclear receptors except for two DAX1 and SHP which both lack the ability to directly associate with DNA [11]. Typically composed of between 70-80 amino acids the pivotal DBD is structurally known to be helix-loop-helix with the two alpha helices packed perpendicular to the center of the polypeptides [12, 6]. The DBD possess two zinc ions chelated by cysteine (histidine residues are common as well) residues commonly referred to as zinc fingers at the start of each helix. Within and between these 2 cysteine rich zinc fingers are several sequence elements termed P, D, T, and A boxes [7]. These elements play crucial roles in response element specificity, dimerization interface and interaction with DNA backbone. A modification of amino acid sequences within this region, specifically the P-box can result in interconversion of nuclear receptor response element recognition [11].

3.2.3 Ligand binding domain

The LBD or ligand binding domain is the staple of the nuclear receptor; located within the E-region which additionally houses a ligand-dependent activation factor termed AF-2 [13]. The LBD of nuclear receptors demonstrate a large amount of conservation between receptors and families and still maintains enough divergence to remain ligand selective. Multiple crystal structures of this region have been solved both independent and in the presence of agonist, antagonist and peptides. The LBD structurally consist of 11-13 helices arranged in three layers [14]. Within these helices exist a ligand binding pocket (LBP) which is guarded by a twelfth

helix [13]. This helix twelve (H12) forms a lid over the LBP and contains residues crucial for the function of AF-2, co-activator or co-repressor recruitment [13,15]. Structural insight of the LBP has operated as the centerpiece for nuclear receptor modulation research. Broadly the LBP can be characterized as a cavity highlighted by nonpolar amino acid residues while possessing few polar residues. Structurally the LBP is divergent across receptors and within receptor families varying in shape, size, volume and residues. The disparity within the LBP provides the distinguishing feature that allows differentiation between ligands.

3.2.4 Activation transcription factor (AF-2)

AF-2 is a well conserved ligand dependent activation factor located on the surface of the LBD [16]. The ability of AF-2 to play a role in co-activator or co-repressor relationships with the receptor has caused this factor to garner a significant amount of attention for nuclear receptor modulation and signaling.

3.3 Nuclear Receptor Domain Function

3.3.1 Activation Transcription Factor-1 (AF-1)

This ligand-independent activation factor regulates gene transcription in co-operation with AF-2 to promote maximal gene expression.

3.3.2 DNA Binding Domain

The DBD serves to read the HRE and bind selective response elements upstream of the target gene predicated on the class of receptor, therefore receptors can be distinguished by specific DNA sequences. The DBD also dictates dimerization, receptors can function as homodimers,

heterodimers or monomers. Several advantages to the binding site arise from dimerization including affinity, specificity and diversity.

3.3.3 Hormone Response Elements

The hormone response elements reside in the promoter gene and binds to the hormone receptor via the DBD this serves to allow transcription of genes under the influence of the hormone to be regulated by means of the receptor.

3.3.4 Hinge

The hinge plays an important role in conformational changes by not only connecting the DBD to the LBD but in a manner as to provide flexibility. The importance of this act is to ensure conformational changes can occur without the presence of steric hindrance. In addition, the hinge also serves to house the nuclear localization signal (NLS) which facilitates the receptors uptake into the nucleus.

3.3.5 Ligand Binding Domain

While the ligand binding domain contributes to the dimerization interface shared by the DBD, it predominately houses several important modules critical for ligand mediated response; of which none more important than the ligand binding pocket. The ligand binding pocket serves to bind and transcribe the cryptic information integrated within the ligand while providing protection from the external cellular environment in a solvent free surrounding.

3.3.6 Activation Transcription Factor-2 (AF-2)

The primary role of AF-2 is the recruitment or release of co-activators or co-repressors respectively. The course of action embarked upon by AF-2 is a consequence of the presence of a ligand thus making the transcription factor ligand dependent.

3.4 Nuclear Receptor Ligands

Nuclear receptor ligands both natural and synthetic are diverse in structure and function across the superfamily of receptors as illustrated in Fig. 3.2. The very nature of the ligands ability to bind within the nuclear receptor demands that ligands are defined by certain criteria. Typically, small and mainly hydrophobic molecules many of which contain polar functionality affixed at the ends of the molecule to form hydrogen bonds or electrostatic interactions to increase stability. Irrespective of the class of nuclear receptor, ligands can be grouped into two headings natural or synthetic. Natural ligands are those derived from biosynthetic pathways (lipids, PG's, hormones) including derivatives or metabolites of the products biosynthetic pathway. In opposition synthetic ligands are lab engineered, often using endogenous ligands as a basis for design or starting point. Whether natural or synthetic, ligands predominately are divided into classification predicated upon the response generated from the receptor.

The wide distribution of nuclear receptors within the body makes ligands that are both receptor and tissue selective often referred to as selective nuclear receptor modulators of increased interest. Ultimately the observed physiological response is a direct result of the encoded message transcribed by the ligand. Ligands are determined to be agonist or antagonist both of which headline several sub classifications as descriptors.

3.4.1. Agonist

Agonist with respect to nuclear receptors by description are compounds that bind and allosterically promote dissociation of co-repressor proteins and recruitment of co-activator accessory proteins.

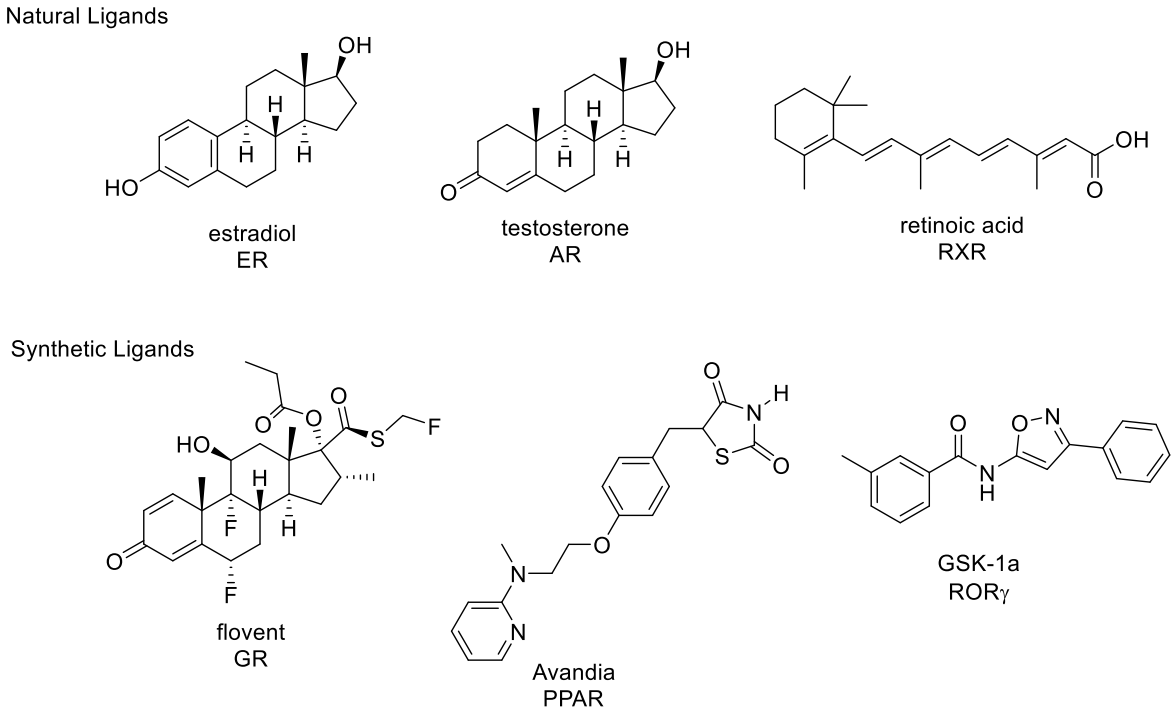


Figure 3.2 Natural and synthetic ligands for various nuclear receptors. Diverse agonistic ligands for various nuclear receptors are shown both natural ligands and synthetic.

Under the general classification of agonist are additional titles offering more refinement. Agonist ligands can elicit maximal response (full agonist), response greater than that of the endogenous ligand (super agonist) or response lower than that of maximal (partial agonist). Computational calculations have demonstrated many nuclear receptor agonist ligands fall within a molecular volume of 250-350 A^3[16]. In general molecules capable of promoting agonistic behavior within the receptor are predicted to be hydrophobic, rigid in skeleton and possessing polar groups on the ends of the molecule. In addition to natural agonist, synthetic agonist for nuclear receptors exist and have done very well commercially such as Targretin (Bexarotene) an RXR agonist approved by the FDA for treatment of cutaneous T-cell lymphoma.

3.4.2. Antagonist

Antagonist bind to the nuclear receptor within the ligand binding domain but promote detainment of co-repressor proteins thus suspending the transcription pathway. Ligands that can be classified as antagonist upon bonding can further be described as affinity reliant for retention (competitive antagonist), affecting the magnitude of the response capability (noncompetitive antagonist), allosterically influenced by a pre-exposed agonist (uncompetitive antagonist) or possessing zero efficacy for activation (silent antagonist).

Generally, antagonist have the skeleton of the agonist however the antagonist possesses a larger scaffold extending from the molecule commonly arising from the center of the skeleton with the intent to disrupt helix packing (Ex: Diethylstilbestrol vs Tamoxifen shown in Fig. 3.3). Analogous to agonist, synthetic antagonist compounds have achieved a considerable amount of commercial success with the most notable being Tamoxifen used in the treatment of breast cancer.

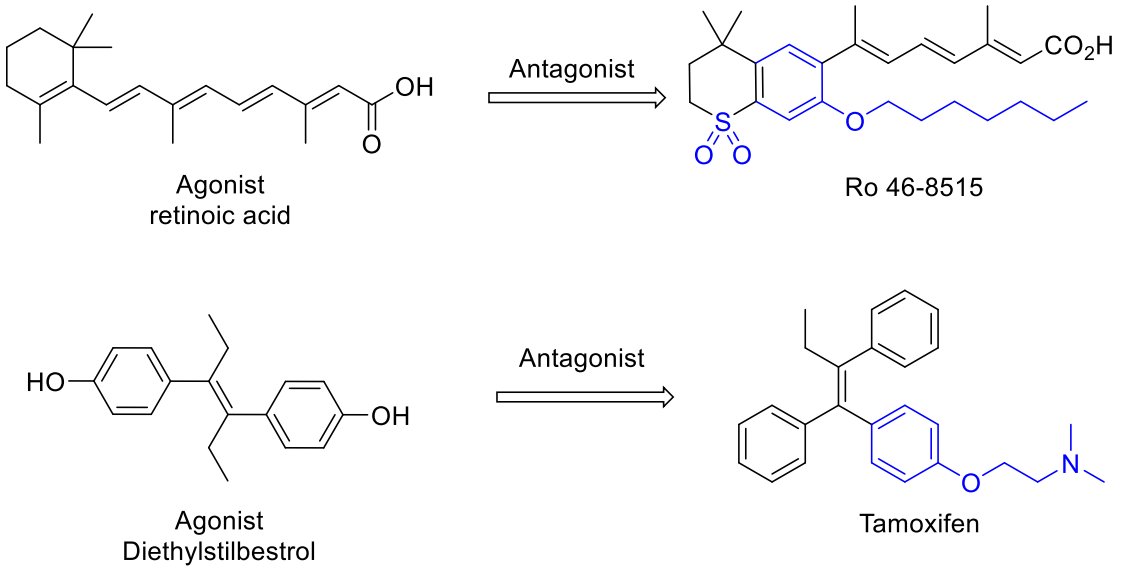


Figure 3.3 Shared Skeleton of agonist conversion to antagonist. Commonly the skeleton for agonist and antagonist are the same with more bulk (highlighted in blue) added to the antagonist.

3.5 Mechanism of Ligand Mediated Receptors

The turn on of gene transcription via nuclear receptor modulation involves a sequence of events composed of complex association and auxiliary co-factors. The receptor resides in two states an inactive state (apo) which exist in the absence of a ligand or an active state (holo) which is ligand bound and thus commences a cascade. For mechanistic action of the nuclear receptor to occur, the receptor must be able to bind to a ligand then recognize HRE near or in the target gene and finally alter transcription rate of the target promoter. It is only once the ability to satisfy these criteria can be meet that the mediated mechanism can begin.

The binding of the ligand is the first pivotal event that must occur for mediated transcription to occur. Once bound the ligand is completely engulfed within the LBP and contacts with amino acid residues within this highly hydrophobic core are formed with the ligand. A ligand induced conformation change also occurs of the LBD, repositioning helix 12 this alteration predominately exists in steroid receptors. The new orientation of helix 12 forms a seal over the LBP and is direct consequence of the bound ligand. The movement of helix 12 has been depicted in two models the mouse trap model and the dynamic stabilization model. This induced conformational change generates a transcriptionally competent AF-2 and allows for the continuation of the mechanistic scheme.

Association of coregulatory complexes with the receptor is the next crucial occurrence, ligand binding induces an allosteric effect in AF-2 signaling recruitment of auxiliary proteins. Coregulatory proteins serve one of two purposes either promotion of activation (co-activators) or repression (co-repression) each occurring in response to the presence or absence of a ligand. In the presence of an activating ligand (agonist) co-activators such as steroid receptor coactivator (SRC-1) are recruited. Residing within this peptide is a LxxLL (L=leucine, x = any amino acid

residue) motif otherwise known as a nuclear box that interacts directly with the LBD of the receptor. This coactivator-receptor interaction occurs within a hydrophobic groove involving residues in helices 3 and 12, lysine and glutamic acid respectively hydrogen bonding with carbonyls in the backbone of the nuclear receptor box [7]. These co-activator, serve to bind acetylated histones thus weakening histone-DNA interaction resultant in relaxed chromatin from the promoter gene. After this exchange of co-regulators, RNA polymerase II is recruited and mRNA transcription is initiated. In the absence of a ligand co-repressors are bound, that have yet to undergo disassociation from the receptor the most commonly of which have been studied are nuclear corepressor (NCoR) and silencing mediator for retinoid and thyroid receptor (SMRT). Similar peptide motifs found in the nuclear receptor boxes are also located within the corepressors however interactions do not involve helix 12. These corepressors serve to recruit histone deacetylase (HDAC) which results in condensation of chromatin above the promoter gene ultimately repressing transcription. Subsequently binding of an antagonist involves retention of corepressors, where upon conformational change places helix 12 in an orientation that prevents recruitment of co-activators.

The ligand-ligand binding domain relationship can be viewed as input-output system wherein the ligand (input) binds within the LBD, induces allosteric changes at the receptors surfaces and the information encoded within the small molecule is communicated to the intracellular environment (output) [15]. The margin for error within the signaling communication process causes nuclear receptors to be implicated in a vast amount of diseases.

3.5.1. Apo

Not many crystal structures have been generated for unliganded receptor ligand binding domain since the addition of the ligand generally adds stability. Under the guides of the mouse trap

model when a receptor is void of a ligand helix 12 is far away from the body of the LBD. Contrary to the mouse trap model, the dynamic stabilization model indicates that even in the absence of a ligand helix 12 is close to the body of the receptor. The dynamic stabilization model even states that in some cases helix12 proximity is so close that it forms a lid over the LBD during an inactive state, however the helix is not fixed and labeled as more molten. While it appears, disunity may arise with respect to the positioning of helix 12 during the Apo state consensus is the unliganded receptor allows more fluidity in the LBD and less restriction to the surrounding helices.

3.5.2. Halo

Much of crystal structures solved for nuclear receptor ligand binding domains exist within the Halo or bound state. As described with the mouse trap model once in the bonded state helix 12 moves proximal to the ligand however this model has only proven true for two receptors RAR and RXR LBD [17]. The dynamic stabilization model describes the introduction of ligand as lowering folding energy globally for the receptor. This thereby reduces conformational dynamics and fixies helix 12 in a stable conformation. Under the scope of the dynamic stabilization model receptors that are constitutively active can be unliganded and stable on account of helix 12 already having the most stable conformation.

3.6 Nuclear Receptors Role in Diseases

Nuclear Receptors have been linked in a variety of roles to various diseases [18]. The correlational between receptor under expression and occurrence of disease has been implied with vitamin D receptor and its link to type I diabetes mellitus [19], estrogen receptor β and its link to

colon and prostate cancers [20]. Oppositely overexpression has been observed with estrogen receptor α in mammary gland estrogen receptor positive tumors [21]. In addition to varying expression levels linking diseases to nuclear receptors, process regulation controlled by the receptor can trigger responses that aid in the treatment of diseases one such example is regulation of apolipoprotein E (ApoE) and its lipid transporters via liver X receptor (LXR) aids in amyloid β clearance in the treatment of Alzheimers disease [22]. Several studies have shown that deregulation of nuclear receptors can play pivotal roles in specific human diseases [18,19]. Nuclear receptors including the estrogen receptor (ER), peroxisome proliferator-activated receptor gamma (PPARγ) and retinoid x receptor (RXRα) have emerged to the forefront of investigation for implications in diseases ranging from neurological, metabolic and cancers.

3.6.1 Estrogen Receptor

The role of estrogen receptor in human disease received wide spread attention with its link to breast cancer, however the receptor has been implicated in other types of cancers. Associated with ovarian, colon and suggestive in prostate cancers estrogen receptor is believed to play varying roles in different types of cancers. In some cancers (ovarian) high level of estrogen expression is a disservice with cancer cell growth supported by the presence of an agonist. In other forms of cancer, a more protective role is initiated where agonist is needed due to low expression levels of estrogen receptor.

In addition to cancers estrogen receptor has been linked to neurological as well as cardiovascular diseases. Evidence has suggested within cardiovascular disease that estrogen receptors have both a direct and indirect role lowering the risk for coronary disease. One of the direct role estrogen receptor families play is via its ability to be stimulated to trigger the rapid release of nitric oxide

within cardiovascular smooth muscle. The significance of this release is that nitric oxide relaxes vascular smooth muscle which in turn inhibits platelet aggregation.

3.6.2 Peroxisome Proliferator-Activated Receptor gamma

The peroxisome proliferator-activated receptor and its sub-types have demonstrated the ability to mediate inflammatory responses via ligand regulation [23,24]. Neurodegenerative diseases all induce potent inflammation in the brain as the immune systems response to disease-initiated agitation. Control over inflammatory response can be governed by a diverse cluster of immune mediators including pro-inflammatory cytokines. In addition to inflammation several neuro-diseases such as Huntington and Alzheimer show high levels of protein (amyloid) build up. Other degenerative diseases, Parkinson for example show reduced levels of specific co-activators mainly PCG-1α (principal cofactor). PPAR and in addition to other families of receptors are highly expressed in the brain and all cell types in the central nervous system. A significant number of studies have exemplified the ability of PPAR agonist to suppress cerebral inflammation and reduction of amyloid levels in mice with Alzheimer's. In addition to inflammatory regulation the PPARγ subtype action is bound to expression of PCG-1α, therefore binding of agonist stimulate expression. PCG-1α deficiencies have also been shown to be culprit in other mechanistic pathways such as insulin sensitivity, energy production and neuronal viability. Metabolic diseases including but not limited to obesity, type II diabetes and hypertension have all been linked to cellular factors that are regulated by nuclear receptors [25].

3.7 Nuclear Receptor Role as Pharmaceutical Targets

Drug design and discovery has chiefly been fueled by two systematic approaches: phenotypic screening and target-based design, with the latter affirming its position by the end of the mid-20th

century. In target-based approach, a selected biochemical mechanism is proposed to be implicated in a disease(s) and modulation of biochemical factors within these mechanisms serve as targets for drug discovery. Classical target-based approach centers around deconstructing the biochemical pathways to develop a strengthened mechanistic derived targeted approach. The benefit of this type of approach is often the biochemical factor (**target**) can be incorporated in screening thereby increasing efficacy of lead discovery process while minimizing discovery time [26]. Additionally, target inclusion in screening provides the opportunity to gain insight about the interaction between the ligand and the biochemical factor of interest. Critical information can be gained through understanding these relationships such as hydrophobicity, donor-acceptor roles and ligand restrictions, ultimately leading to more rationale screening. Target based derived approaches resulted in Food and Drug Administration (FDA) approval of 78 out of 113 first in class drugs (1999-2008) [27] facilitating the emergence of target-based drug discovery approach as an applicable tool for new drug identification.

Nuclear receptors are a class of proteins that have often been subjected to target-based approach for the discovery of potential new ligands for these receptors. As ligand activated transcription factors, these receptors play an integral part in several metabolic and developmental processes; thus, have been implicated in several diseases, ranging from atherosclerosis to cancer. To date there are 48 human nuclear receptors that have been identified, each of which can regulate expression of genes under its control thus linking receptors to key physiological responses [3]. The potential for error within the ligand-initiated communication between the nuclear receptor and the signaling process for change of cellular function solidifies nuclear receptors role in diseases.

Modulation or inhibition of these biological pathways to turn on/off key physiological responses for treatment of disease makes nuclear receptors “pharmaceutically significant”.

Nuclear receptors ability to act in concert with a host of cellular factors in response to introduction by a ligand provides a stage for their role in drug discovery. Nuclear receptor signaling targets accounted for 107 new FDA approved drugs by the end of 2013, making them the third most targeted and fastest-growing class of protein targets, with almost two small molecules approved per year for these receptors [28]. The growth of interest in ligand generation has not gone unnoticed by the pharmaceutical industry with 33% of nuclear receptors being targets [29] estimates suggest that nuclear receptor ligands constitute 10-20 % of worldwide pharmaceuticals [30]. Not only have nuclear receptor ligands claimed stake in the pharmaceutical industry but have exceled, as of 2003 the top 200 most prescribed drugs included 34 that targeted nuclear receptors [31]. Pharmaceutical industry involvement in receptor therapeutic ligand development has resulted in [32] several commercially successful compounds. As illustrated in figure 3.4 ligands such as Tamoxifen for treating breast cancer *via* the estrogen receptor, Flovent, a corticosteroid used to treat asthma by selectively targeting the glucocorticoid receptor, Casodex, an antagonist for the androgen receptor in the treatment of prostate cancer and Avandia, the antidiabetic drug which binds the peroxisome proliferator-activated receptor successfully impacted pharmaceutical sales. A developed understanding of how ligands turn on or off transcription has aided in the contribution of commercially successful ligands.

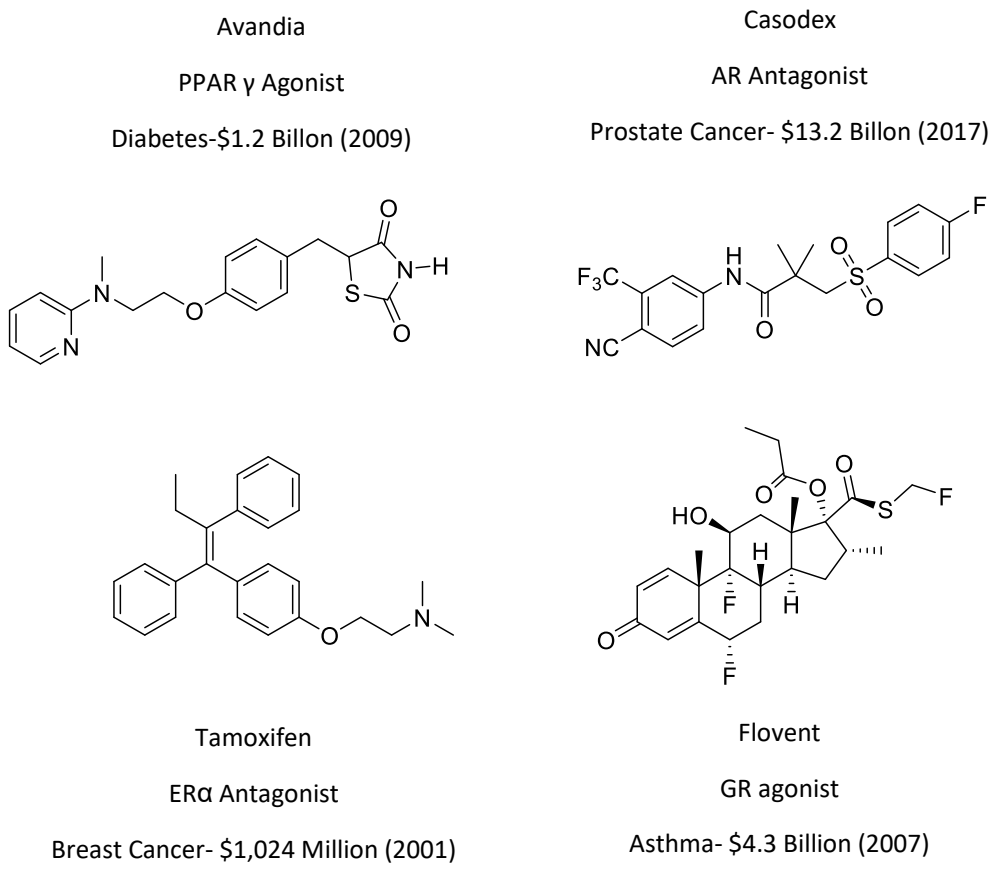


Figure 3.4 Commercially available nuclear receptor ligands. The above ligands are part of the worldwide pharmaceutical market with large sales in the treatment of various diseases and conditions.

3.8 References

1. Germain, P; Staels, B; Dacquet, C; Spedding, M; Laudet, V. *Pharmacol Rev*., **2006**, 58,685–704

2. Gustafsson, J. *Journal of Steroid Biochemistry & Molecular Biology*. **2016**, 157,3-6

3. Maglich, J.M; Sluder, A.E.; Willson, T.M.; Moore, J.T. *Am. J. PharmacoGenomics*. **2003**, 3, 345–353.

4. Noy, N. *Biochemistry*.**2007**,46, 47, 13461-13467.

5. Caboni, L.; Lloyd, D. *Medicinal Research Reviews*, **2013**, 33, 5, 1081–1118.

6. Katzenellenbogen, J.; Katzenellenbogen, B. *Chemistry & Biology* **1996**, 3, 7, 529-536.

7. Germain, P.; Altucci, L.; Bourguet, W.; Rochette-Egly, C.; Gronemeyer, H. Pure and Applied Chemistry, 2003, 75, 1619-1664.

8. M. Fu et al. *Mol. Cell Biol*. **2002**, 22, 3373–3388.

9. C. Wang et al. *J. Biol. Chem*. **2001,** 276, 18375–18383.

10. Bourguet, W.; Germain,P.; Gronemeyer,H. *Trends Pharmacol. Sci*. **2000**, 21, 381-388.

11. Gronemeyer, H etal. *Nature Reviews Drug Discovery*, **2004**, 3, 951.

12. Helsen, C. etal. *Mol Cell Endocrinol.* **2012**, 348(2), 411-7.

13. De Lera, A.; Bourguet, W.; Altucci, L.; Gronemeyer, H. *Nature Reviews Drug Discovery*, **2007,** 6, 811-820.

14. Jordan, V. *Journal of Medicinal Chemistry*, **2003**, 46 (6), 883-908.

15. Nicholes, C.; Liu, K.; Lin, C.Y. *Journal of Cellular Biochemistry*, 2013, 114, 2203-2208.

16. Sladek , F. *Mol Cell Endocrinol*, **2011**, 334, 3–13.

17. Germain, P. *Pure and Applied Chemistry*, **2003**, 75, 1619-1664.

18. Wagner, M.; Zollner, G.; Trauner, M. *Hepatology* **2011**,53, 3, 1023-1034.

19. Levi, M. *Biochimica et Biophysica Acta* **2011**, 1812, 1061-1067.

20. Deroo, Bonnie.J.; Korach, Kenneth S. *Journal of Clinical Investigation*, **2006**, *116*, 561.

21. Mills, J., Rutkovsky, A., Giordano, A. *Current Opinion in Pharmacology*, **2018,** 41:59–65

22. Skerrett, R,; Malm, T.; Landreth, G. *Neurobiology of Disease*. **2014**, 72, 104-116.

23.Tyagi, S. etal. *J. Adv. Pharm. Tech. Res.* **2011** 2, 4, 236-240.

24. Sandeep, T., etal. *J. Adv. Pharm. Tech. Res.* **2011**,4, 236-241.

25. Schulman, I. *Advanced Drug Delivery Reviews.* **2010**, 62, 1307-1315.

26. Brown, D. *Drug Discovery Today*. **2007**, 12, 1007-1012.

27. Wallace, O.; Richardson, T.; Dodge, J. *Current Topics in Medicinal Chemistry*, **2003**, 3, 1663-1680.

28. Kinch, M.; Hoyer, D; Patridge, E.; Plummer, M. *Drug Discovery Today*. **2015**, 20, 784.

29. Fraydoon R.; Pengxiang H.; Vikas, C.; Sepideh, K. *J Mol Endocrinol*, **2013**, 51.

30. Fraydoon R.; Pengxiang H; Vikas C; Sepideh K. *J Mol Endocrinol*, **2013**, 51.

31. http://www.rxlist.com/top200.htm

32. Ottow E & Weinmann H. pp 1–23. Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA. (doi: 10.1002/9783527623297.ch1)

CHAPTER 4

SYNTHESIS OF GFP CHROMOPHORE INSPIRIED LIGANDS AS ESTROGEN RECEPTOR ALPHA AGONIST

4.1 Estrogen Receptor

Estrogen receptor (ER) resides under the classification of steroidal nuclear receptors and exist as two isoforms in mammals, estrogen receptor α (ERα) and estrogen receptor β (ERβ) [1]. Discovered in the 1960's and cloned in the 1980's it was not until after the discovery of a second subtype that the original receptor was renamed estrogen receptor α [2,3]. Like most nuclear receptors ERα and ERβ are ligand regulated, with 17β-estradiol (E2) being its endogenous ligand. It is worth mentioning that receptors similar in nature and function of the estrogen receptor have been discovered termed estrogen related receptors (ERR) however they do not bind estradiol (E2) and no known endogenous ligand exist. [4, 5]. Due to its subcellular mode of action ER is labeled as a Type 1 receptor. In the absence of a ligand ER is cytosolic anchored by chaperone proteins commonly heat shock proteins (HSP). Ligand binding initiates HSP release followed by dimerization and translocation into the nucleus [6]. Once in the nucleus association with transcriptional coactivators that facilitate binding and activation of target genes occurs. [7]. The estrogen receptor family is among nuclear receptors that have warranted considerable attention.

4.1.1 Estrogen Receptor Alpha

While both estrogen receptor subtypes can be considered pharmaceutically significant ERα classically named NR3A1 is a vested therapeutically targeted receptor in contraception, fertility enhancement, hormonal therapy, and breast cancer therapies. Consisting of 595 amino acid

residues ERα is predominately composed of alpha helixes with less than 3% beta sheets and the rest being turns and loops.

Estrogen receptor α mRNA has wide spread distribution within various sites in the body such as mammary gland, ovary, testes and sections on the brain [8]. As illustrated in table 4.1 expression can be observed in high to moderate even low levels within specific tissues and organs. In addition to being located within the cardiovascular and central nervous system the uterus and pituitary gland express ERα however only after maturity of the tissue has been reached [9]

Table 4.1 Estrogen receptor α tissue distribution and expression levels. The above table illustrates the varying levels of ERα expression in different tissues.

Tissue	ERα Expression Level
Reproductive (testies, ovaries)	high
Prostate, bladder and liver	moderate
Epididymis, pituitary gland, thymus	low

4.1.2 Ligand Binding Pocket

The crystal structure of ERα-LBD complexed with 17β-estradiol gave a large amount of insight into the relationship between the ligand binding pocket (LBP) and ligands [9, 10]. The LBP is formed by residues from helices 3,6,8,11,12 as well as the loop region between helix 7 and 8. With a volume size of $450A^{3,}$ the LBP is completely partitioned from the external environment and generates a fairly large portion of the hydrophobic core of the LBD as shown in figure 4.1. Full agonistic behavior is believed to be achieved when the ligand is encapsulated into the hydrophobic binding pocket and H12 can seal the pocket [11, 12, 13]. The importance of conservation of specific residues, spatial availability and flexibility within the pocket has been extensively studied [13,14] with results suggestive that ligand volume (size, shape), hydrogen

bonding network and ligand molecular topology are crucial criteria for successful activation. Estrogen receptor α's wide distribution in human tissue and roles in many physiological systems links it to diseases such as osteoporosis, cardiovascular, obesity, Alzheimer, autoimmune and cancer (breast & ovarian) [15, 16, 17]. Thereby continuing the interest in development of ligands that bind ERα.

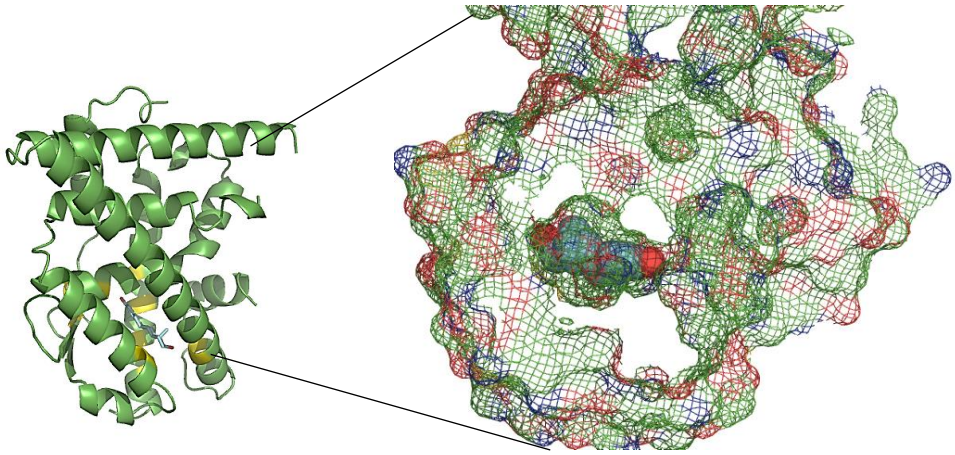


Figure 4.1 Estrogen Receptor α complexed with estradiol. Secondary structure of ERα with estradiol (E2 in green), inset: ligand binding pocket residues complexed with E2 exemplifying the depth of pocket in the protein.

4.1.3 Ligands

Successful modulation of ERα has been demonstrated across a broad class of ligands including both steroidal and non-steroidal. A comprehensive investigation into the two categories of ligands demonstrate agonist and antagonist for ERα that are synthetic, non-synthetic and secondary metabolites as illustrated in figure 4.2 [18]. ERα ligands can be divided into several classes including natural steroidal, synthetic non-steroidal and plant secondary metabolites [19]. The ability of ERα to bind such a diverse set of compounds would suggest promiscuity in binding, however despite the origins of these ligands in relation to diversity in species, there are commonalties amongst the diversity. Several studies have been published investigating common

core features of estrogen receptor ligands. Of those studies Katzenellenbogen et al. identified a generic model estrogen receptor pharmacophore [14] as illustrated in figure 4.3.

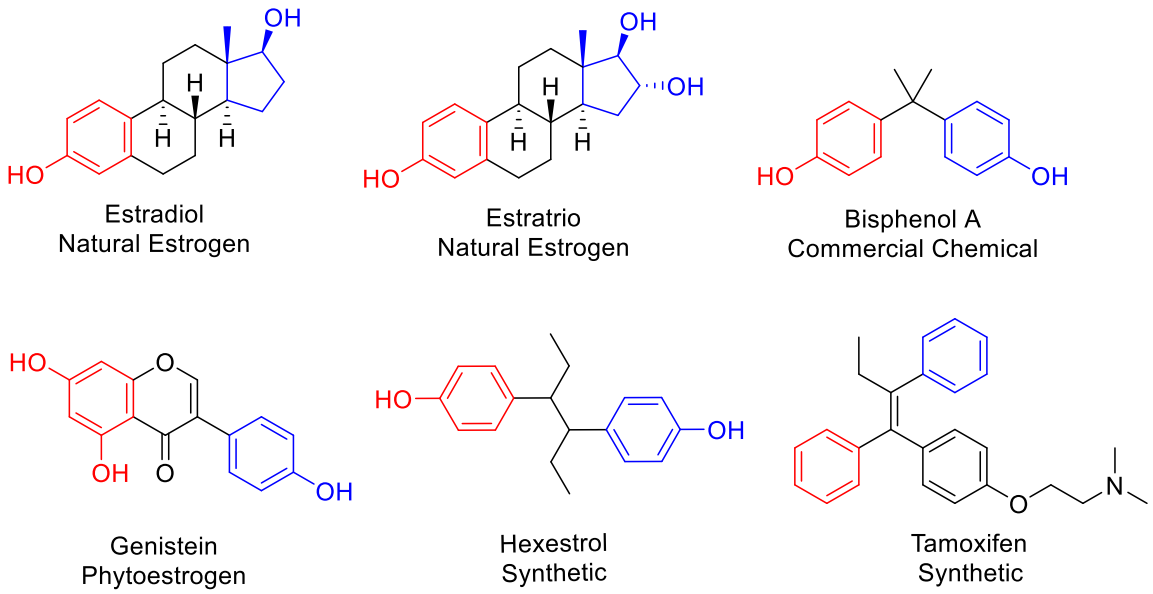


Figure 4.2 Known estrogen receptor ligands. Identified estrogen receptor ligands with requirements highlighted (blue & red) in relation to established Katzenellenbogen ERα pharmacophore.

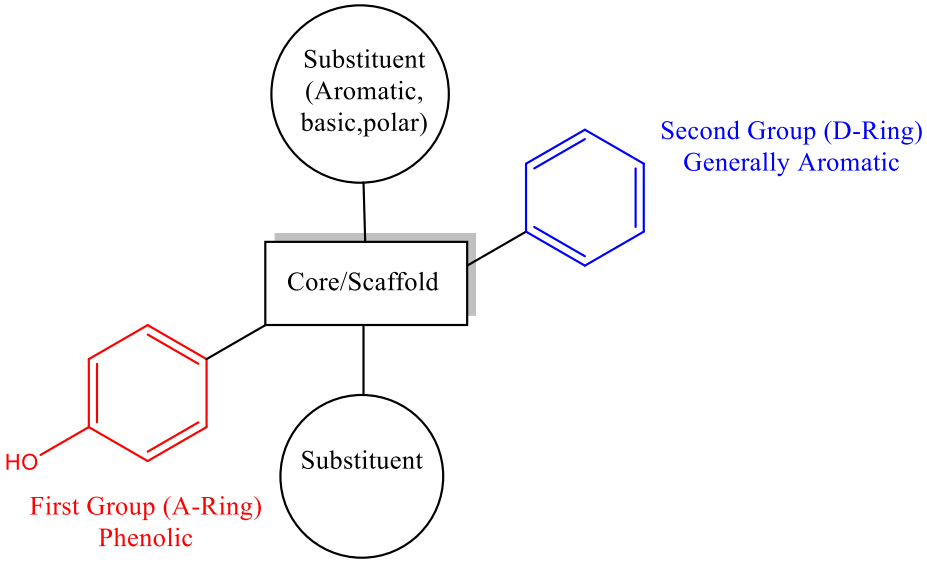


Figure 4.3 Estrogen receptor pharmacophore. The above figure illustrates the estrogen receptor non-selective pharmacophore established by Katzenellenbogen with distinguishing features highlighted in red and blue corresponding to figure 4.2.

Lengthy studies into known estrogen receptor modulators and their potential interactions within the receptor ligand binding pocket enabled Katzenellenbogen to unravel features that are exchangeable from those that are indispensable. Reviews have been compiled in which various categories of synthetic compounds and natural product analogs support this model [20, 21.]. While estradiol, ERα endogenous ligand skeleton is composed of an ABCD ring system, this proposed pharmacophore embraces estrogen receptor modulators that only contain ACD ring systems. The commercial success of ERα modulators has exacerbated the development of additional ligands for physiological perturbation.

4.2 Estrogen Receptor Ligand Design Rationale

Guided by investigation into the photo physics of the green fluorescent protein chromophore our lab extensively studied the arylmethyleneimidazolone (AMI) core [22-26]. Generated via a post translational autocatalytic cyclization followed by autoxidation involving specific residues within the protein this core is composed of key features exemplified in known estrogen receptor modulators. In addition to notable features, the AMI skeleton is amenable to being reminiscent of the ACD ring scaffold proposed by Katzenellenbogen [14]. Estrogen receptor's endogenous ligand estradiol possess a ABCD ring structure however an established precedence exists for utilizing a ACD ring system. Also termed B-ring seco-steroids and B ring seco-estradiols initial reports of ACD systems surfaced in 1959 by Novello [27]. Early estrogenic activity was studied by Bindra, Neyyarapally and co-workers in ACD systems [28, 29]. Katzenellenbogen reported of ACD systems inspired by genistein with high potency and isotype selectivity [30]. The ACD system boast principal advantages over the classical ABCD ring system primarily the easy of functionalization in the A and D systems. The established foundation and success of the ACD

system with binding and selectivity in the estrogen receptor prompts discovery for other sources of ACD system inspiration.

Visual inspection of known estrogen receptor α ligands and AMI analogues previously synthesized by Baldridge [31] display significant structural commonalities that can be viewed as required for preservation of criteria necessary for modulation.

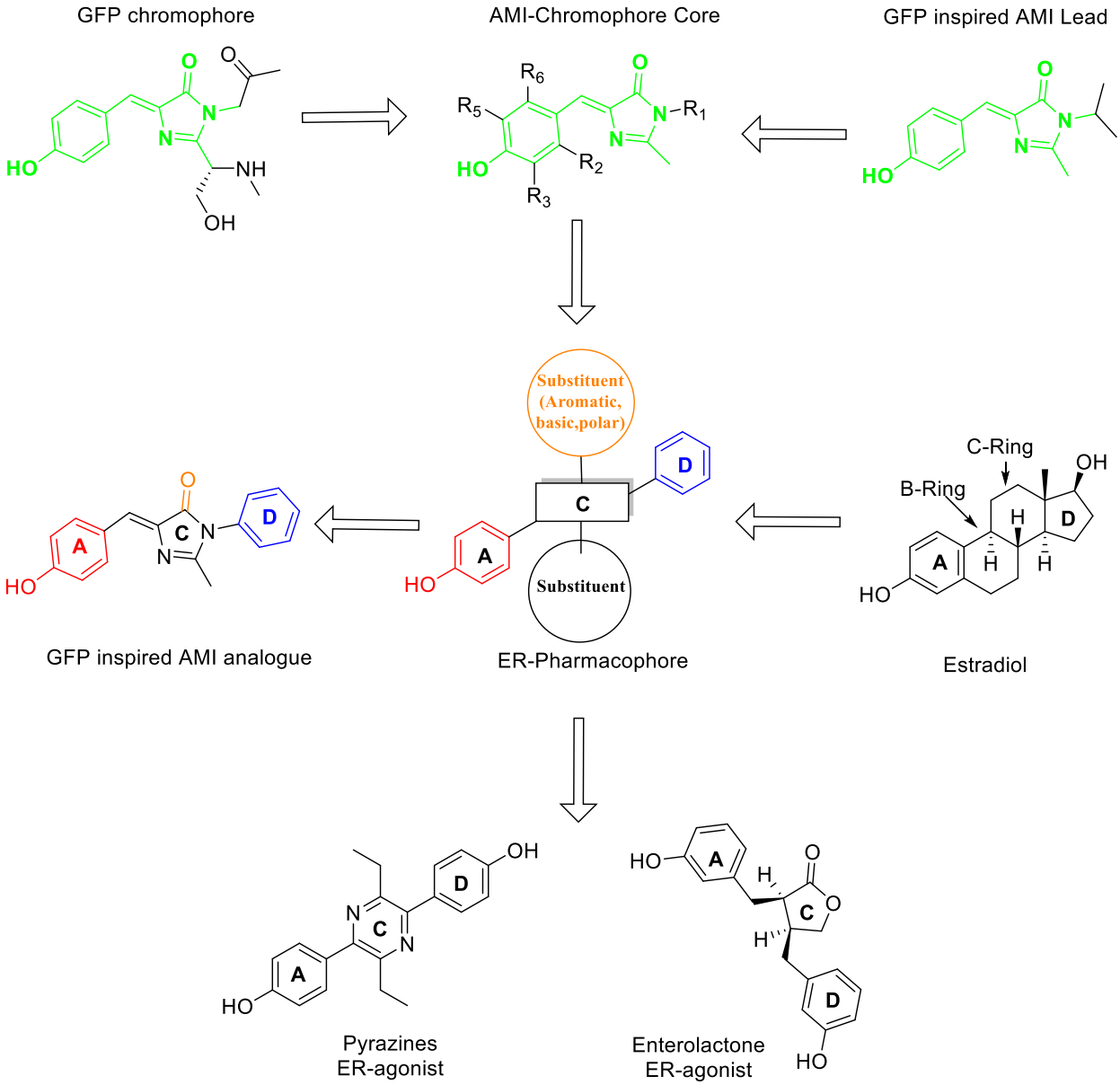


Figure 4.4 Rationale for development of GFP inspired arylenemethylene imidazolone analogue. The above figure highlights the initial GFP core (green) as being amendable to structural changes to satisfy ER pharmacophore requirements colored blue and red. Support is also provided for the use of ACD ring structure with pyrazine and enterolactone example.

Overall the usage of AMI chromophores as ligands for estrogen receptor α activation highlights several advantageous points including providing a skeleton pre-equipped with key features for estrogen receptor activation, a tunable scaffold that may lead to enhanced selectivity here-go minimizing undesired side effects while maintaining facile generation of synthetic analogues. Herein we report a novel class of ERα agonist designed from GFP's chromophore core structure. Simple benzamides have been reported by Caldarelli [32] to demonstrate agonist behavior however to our knowledge this is the first example of GFP-derived class of agonist. Using chemical complementation in yeast and luciferase assay in mammalian cells a library of compounds was screened for agonistic behavior. Additional AMI's were designed and synthesized based on results and a structure activity relationship (SAR) developed for top activators.

Previously defined SAR through the work of Katzenellenbogen and others in designing acceptable ligands for estrogen receptor mandates that consideration be given to certain criteria. With these benchmarks in mind ligands were designed to adhere to the following guidelines:

1) Maintaining the hydrogen bonding network as defined by the ligands relationship with the following residues Glutamine-353, Arginine-394 and Histidine-524.

2) Containing ligand size and volume to requirements in agreement with that of the ligand binding pocket.

3) Maintain molecular topology as defined by polarity of the molecule with specific attention given to the scaffold or core.

4.2.1 Hydrogen Bonding Network

The preservation of hydrogen bonding networks around the Gln-353, Arg-394, His-524 residues were shown a key contributing factor for activity of the ligand [10]. Estradiol has showcased the importance of this network (Figure 4.5) in several aspects.

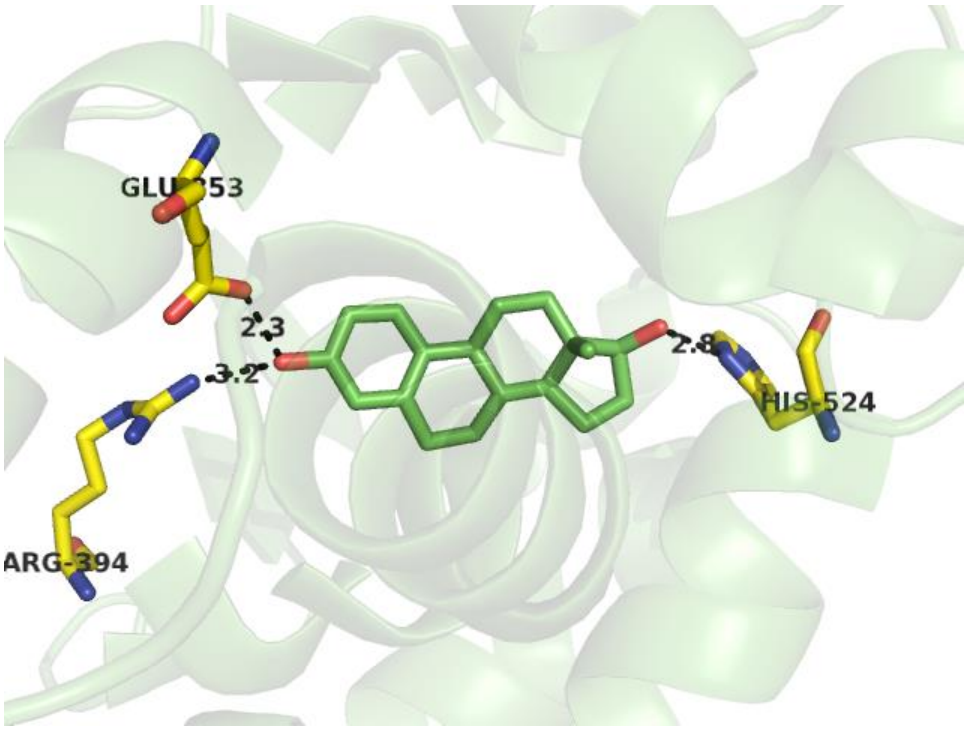


Figure 4.5 Ligand binding pocket of estrogen receptor α. Critical residues highlighted in yellow sticks establish a hydrogen bonding network with estradiol (green). Residues and estradiol contact are under 4A of proximity.

Aside from contributions by possible weak polar interactions from the aromatic ring the hydrogen donor effect of the phenolic moiety (A-ring) with Gln-353 and Arg-394 has been outlined in a series of reviews [14, 33]. Likewise, the ability of histidine or tyrosine to serve as hydrogen donor has defined the role played by His-524 with the 17β position of estradiol [14]. Therefore, to design ligands compounds were generated to explore the contribution from each facet of the hydrogen bonding network. Compounds containing a combination of accessible

(CW28), capped (CW24), sterically hindered (CW25, CW27), or meta/ortho/para substituted (CW29, CW30, CW31, CW32) phenolic moieties were synthesized to evaluate the role of the relationship with Gln-353 and Arg-394. An analogue bearing no hydroxyl substituent (CW33) was included to serve as a control for this series of compounds. In investigating the role of His-524 as a hydrogen bond donor or acceptor the amide was alkylated with a secondary phenolic ring. Attenuations of the hydroxyl group of the newly added phenol generated ortho and para substituted compounds yielding CW36, CW34 respectively. In addition, CW39 was synthesized that methylated the secondary phenol rendering its role to hydrogen bond acceptor exclusively. It is important to note that the primary phenol located on the A-ring on the AMI analogues were held in the para position for consistency with the phenol location in estradiol's A-ring ensuring that any diversity in activation was a direct result of modification of His-524 roles.

4.2.2 Ligand Volume and Size

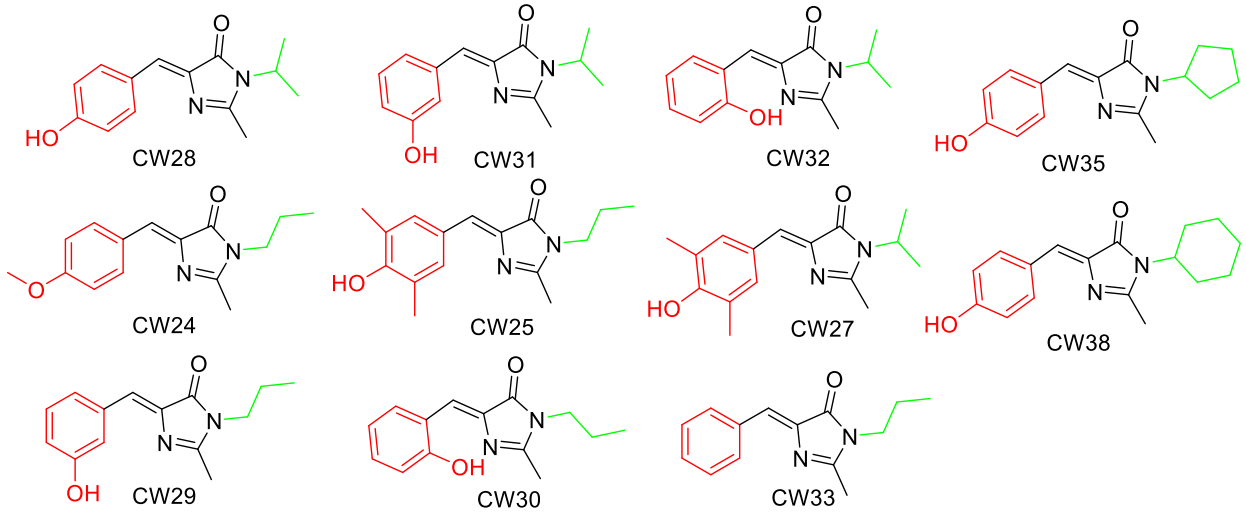
Ligand volume (size and shape) have been linked as co-factors that contribute to modulation with the proper size ligand possessing the ability to induce receptor folding. Katzenellenbogen et al. demonstrated with a series of acyclic amides employing a homobibenzyl backbone that activation is not observed until a ligand of sufficient size was generated [34,35]. Literature has also shown that while estrogen receptor α LBP can accommodate larger ligands highly effective agonistic ligands have molecular volumes significantly smaller than pocket spatial availability including estradiol and diethylstilbestrol at $250A^3$ [19]. Volumetric calculations of the AMI core (unsubstituted) shows a molecular volume of $179.8A^3$, perturbations in volume are rooted in substituent selection therefore substituents were strategically selected as to enhance potency in accordance with literature while not violating the spatial threshold. Small nonpolar moieties such as simple alkyls, as well as polar and basic aryl substituents are implicated in enhanced potency,

for instance n-propyl and isopropyl, substitutions at the 11β (E2) position have been linked to improved binding affinities [2]. To that avail propyl and isopropyl substituents were added to analogues previously constructed to explore hydrogen bonding network. Analogues CW24, CW25, CW29, CW30, CW33 were all alkylated with n-propyl substituents were as CW27, CW28, CW31, CW32, were all alkylated with isopropyl substituents. Cyclic substituents such as cyclohexyl and cyclopentyl rings CW35 and CW38 respectively were added first to increase ligand size and secondly further exploration of isopropyl spatial limitations with regards to improving binding affinities.

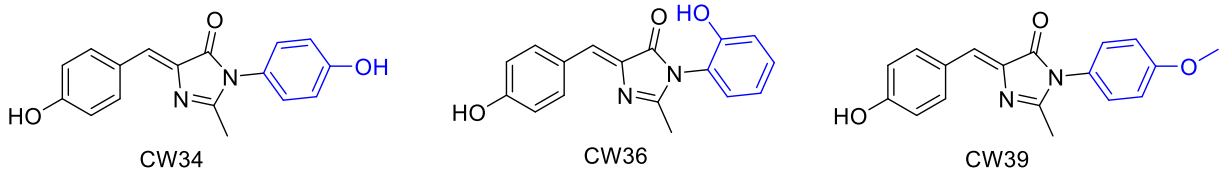
4.2.3. Molecular Topology

Lastly molecular topology i.e. polar surface area seems to be influential in potency of binding affinities. While substitutions may attenuate molecular polarity ultimately the core or scaffold provides an inherent contributor. The incorporation of hetero-cyclic scaffolds to provoke estrogen receptor α modulation is not unprecedented, however studies associate the greater the polarity in the core a greater loss in activity possibly due to a higher desolvation penalty within the LBD [36]. The benzyl methyleneimidzolone is our intrinsic core however analogues were synthesized bearing tetralin (CW1, CW7, CW16) and napthyl (CW59, CW60) substituents attached to the core to reduce total molecule polarity while still generating a ligand of considerable size.

A-Ring Exploration Compounds & Ligand Size-Shape



Role of His 524 Exploration Compounds



Molecular Topology Exploration

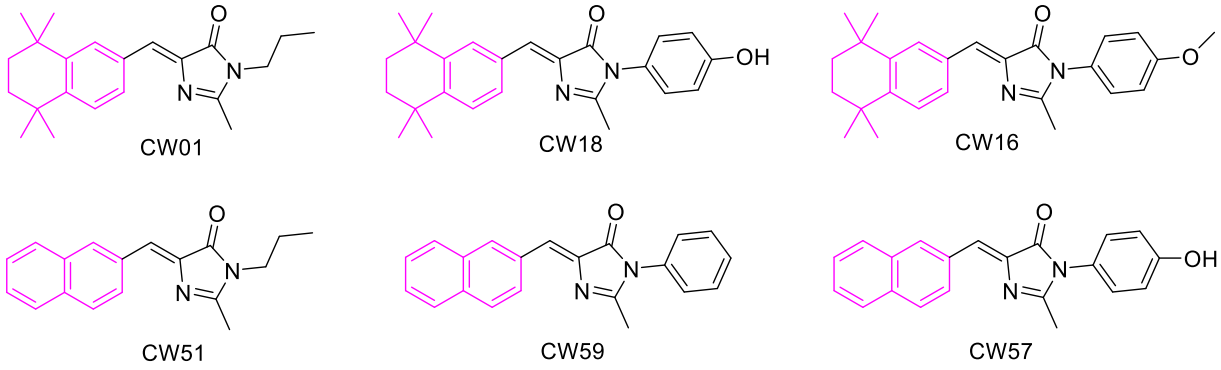


Figure 4.6 Synthesized generation I compounds. Above compounds were synthesized to explore criteria of estrogen receptor α ligand binding pocket residues and this class of compound. Highlighted are A-ring modifications (red), ligand size/shape (green), and molecular topography (violet).

4.3 Synthesis of Arylmethyleneimidazolone

Successful creation of an AMI library is reliant upon a synthetic methodology that allows for facile generation and derivatization of analogues. In addition, methodology must be amenable to diversification of substituents to allow for maximal exploration of the residue contacts within the ligand binding pocket. Dr. Baldridge generated an approach based upon adaptations to methods established by Bazureau that allowed for a more combinatorial route [37]. Baldridge's approach envisioned the substituted arylmethyleneimidazolone (**3**) arising from the 2+3 cycloaddition of an iminoglycine (**2**) methyl ester and a Schiff base (**1**). The Schiff Base was generated by addition of 1.1 equivalent of primary or secondary amine in the presence of an aryl aldehyde and stirred overnight. In cases where both reagents were liquid reaction was run neat, when a solid was present methanol was used as a solvent as illustrated in figure 4.7 below.

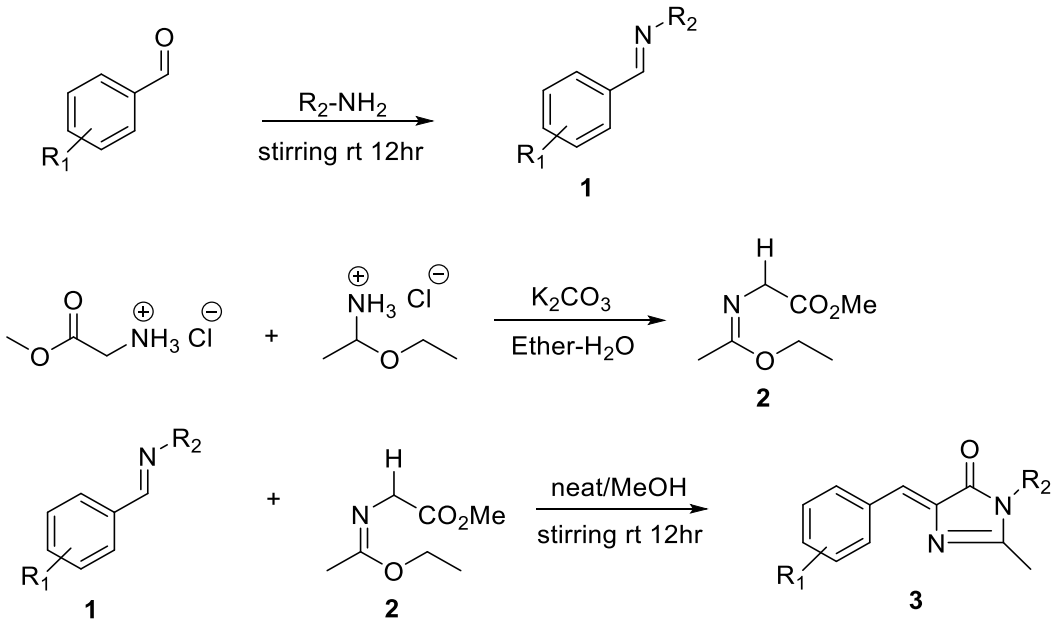


Figure 4.7 Synthesis of Arylmethyleneimidazolones.

Construction of the Schiff Base is the hallmark for diversity in the generation of analogues. The limitation of diversity that can be incorporated on the aryl ring or amide is solely contingent upon the availability of aryl aldehydes and amines, with a broad number of these reagents commercially available as well as synthetically accessible opportunities are vast. In addition to commercially available aldehydes, the tetralin core was generated in three steps from known literature procedure [38,39].

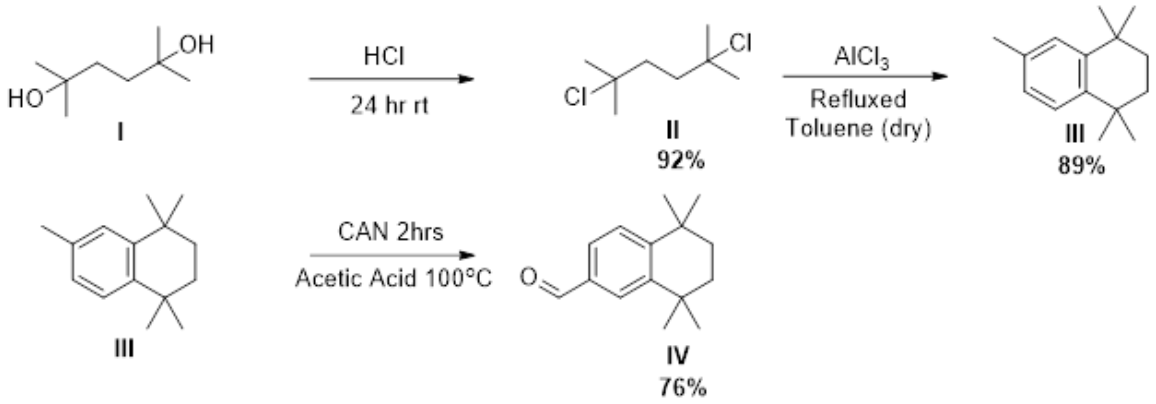
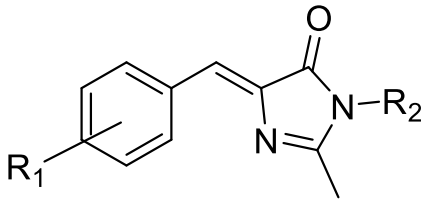


Figure 4.8 Synthesis of tetralin aldehyde

Starting with the commercially available aliphatic diol (I) conversion to the dihalide (II), was carried out by addition of concentrated hydrochloric acid. The substituted aryl ring (III) was generated via Fridel Crafts Alkylation and after oxidation yielded the tetralin based aldehyde (IV) as shown in Figure 4.8. Formation of the iminoglycine methyl ester was carried out by the addition of potassium carbonate, methyl glycimate hydrochloride and ethyl acetimidate hydrochloride in ether with vigorous shaking and frequent venting. Cycloaddition was carried out by addition of 1.0 equivalent of iminoglycine methy ester with 1.0 equivalent of Schiff base and stirred overnight in methanol. Most of reactions precipitated as yellow or tan powders which were filtered, washed with cold methanol and dried in vacuo. Crystallization was achieved for several products at 0^0C for 2-3 days resulting in yellow to orange crystals. A smaller subset of

products was isolated via column chromatography using ethanol/hexane gradient. Utilization of this protocol facilitated synthesis of 72 GFP chromophore inspired compounds with various substitutions. All substituted AMI ligands were synthesized in excellent to modest yields.

Table 4.2 Arylmethyleneimidazolone synthesized ligands



Ligand	R_1	R_2	Ligand	R_1	R_2	Ligand	R_1	R_2
CW1	tetralin	n-propyl	CW35	ρ-hydroxy	cyclopentyl	CW63	2,4-diol	ρ-phenol
CW2	tetralin	n-butyl	CW36	ρ-hydroxy	σ-phenol	CW64	2,4-diol	propyl
CW3	tetralin	n-pentyl	CW38	ρ-hydroxy	cyclohexyl	CW68	2,4-diol	cyclopentyl
CW4	tetralin	n-hexyl	CW39	ρ-hydroxy	ρ-anisole	CW69	2,4-diol	cyclohexyl
CW5	tetralin	isopentane	CW40	ρ-hydroxy	(+/-)phenyl ethyl	CW72	2,3-diol	isopropyl
CW6	tetralin	sec-butyl	CW41	tetralin	n-octane	CW76	σ-hydroxy	cyclohexyl
CW7	tetralin	isopropyl	CW42	tetralin	benzyl	CW77	σ-hydroxy	cyclopentyl
CW8	tetralin	neopentane	CW43	tetralin	n-undecane	CW81	σ-hydroxy	σ-phenol
CW9	tetralin	isobutyl	CW46	tetralin	cyclopentyl	CW82	2,3-diol	σ-phenol
CW13	tetralin	phenylethane	CW47	tetralin	n-cyclohexyl	CW84	ρ-hydroxy	(+/-)-2-propanol
CW15	tetralin	σ-xylene	CW48	napthyl	n-dodecane	CW85	ρ-hydroxy	(R)-2-propanol
CW16	tetralin	anisole	CW49	napthyl	n-octane	CW86	ρ-hydroxy	(S)-2-propanol
CW18	tetralin	ρ-phenol	CW50	napthyl	pentyl	CW90	2,4-diol	(+/-)-2-propanol
CW20	tetralin	4-ethyl-1,2-dimethoxybenzyl	CW51	napthyl	propyl	CW91	2,4-diol	(R)-2-propanol
CW22	tetralin	4-methylpyridine	CW52	3,4-diol	propyl	CW92	2,4-diol	(S)-2-propanol

CW23	tetralin	n-dodecane	CW53	tetralin	(-)phenyl ethyl	CW93	*m*-hydroxy	(+/-)-2-propanol
CW24	ρ-methoxy	n-propyl	CW54	tetralin	(+)phenyl ethyl	CW95	*m*-hydroxy	(S)-2-propanol
CW25	3,5-dimethyl, 4-hydroxy	n-propyl	CW55	ρ-hydroxy	(+)phenyl ethyl	CW96	σ-hydroxy	(+/-)-2-propanol
CW27	3,5-dimethyl, 4-hydroxy	isopropyl	CW56	ρ-hydroxy	(-)phenyl ethyl	CW97	σ-hydroxy	(R)-2-propanol
CW28	ρ-hydroxy	isopropyl	CW57	napthyl	ρ-phenol	CW98	σ-hydroxy	(S)-2-propanol
CW29	*m*-hydroxy	n-propyl	CW59	napthyl	benzyl			
CW30	σ-hydroxy	n-propyl	CW60	napthyl	σ-phenol			
CW31	*m*-hydroxy	isopropyl	CW61	σ-hydroxy	ρ-phenol			
CW32	σ-hydroxy	isopropyl	CW62	2,4-diol	isopropyl			
CW33	H	n-propyl						
CW34	ρ-hydroxy	ρ-phenol						

Table 4.3 Percent yield of ligands synthesized

Compound ID	Yield (%)	Compound ID	Yield (%)	Compound ID	Yield (%)	Compound ID	Yield (%)
CW1	30.4	CW27	10.3	CW49	28.4	CW81	30.2
CW2	69.1	CW28	37.3	CW50	29.9	CW82	67.5
CW3	25.2	CW29	98.4	CW51	17.9	CW84	30.2
CW4	17.1	CW30	53.7	CW52	17.8	CW85	7.2
CW5	47.6	CW31	34.4	CW54	62.1	CW86	17
CW6	37.4	CW32	63.9	CW55	52.6	CW90	37.4
CW7	98.6	CW33	60.4	CW56	54.9	CW91	20.5
CW8	86.6	CW34	34.2	CW57	54.4	CW92	26.1
CW9	95.4	CW35	24.5	CW59	32.9	CW93	28.7
CW13	60	CW36	29.5	CW60	53.2	CW95	18.8
CW15	26.3	CW38	57.9	CW61	55.1	CW96	66.8
CW16	47.6	CW39	33.8	CW62	43.4	CW97	5.1
CW18	17.6	CW40	98.7	CW63	45.6	CW98	52.5
CW20	36.3	CW41	77.6	CW64	48.1		
CW22	24.6	CW42	55.3	CW68	33.1		
CW23	19.4	CW43	49.4	CW69	60.6		
CW24	45.6	CW46	25.3	CW72	39.9		
CW25	29.4	CW47	58.6	CW76	26.6		
CW26	35.2	CW48	42.6	CW77	32.8		

4.4 Arylmethyleneimidazolone Chromophore Properties

The photo physics of GFP, specifically its intense fluorescence, has made it a cornerstone of biological imaging. A major contributor to the strong fluorescence observed from GFP chromophore is the β barrel its enclosed in. In addition to stability and protection from quenching via the external surroundings, residues in the local protein environment assist in excited state proton transfer generating the anionic species of the chromophore ultimately responsible for its

emission. The β-barrel enclosure restricts rotation of the chromophore preventing cis/trans isomerization from occurring which is present in solution with unbound chromophores. In the absence of the protein host synthetic chromophores have drastically different photo physic properties displaying fluorescence several magnitudes less than wild type GFP [38]. Even upon denaturing of wild type GFP fluorescence is weak at best. Reformation of GFP leads to regeneration of the chromophore and restoration of fluorescence [32]. Baldridge et al has demonstrated that enhancement in fluorescence of the synthetic chromophore can be achieved via encapsulation [39]. In addition to enhancement Baldridge could demonstrate proof of this concept with select AMI chromophores and human serum albumin (HSA) [40]. This motivation heightened our belief that AMI chromophores could not only serve as estrogen receptor modulators but also exhibit "turn on" fluorescence. Envisioning the nuclear receptor as the host protein and the LBP as the restricting cavity like that of the β-barrel upon binding the synthetically generated AMI can undergo similar excitation process as wild type GFP thus creating a fluorescently trackable modulator in one package where enhanced fluorescence implies binding.

The spectroscopic properties of the AMI compounds synthesized in this work have been explored and are listed in the Appendix A. All synthesized compounds absorbed in the visible range, with λmax (abs) 345nm-402nm and have fluorescence emissions ranging from 414nm-600nm. Shorter absorption wavelengths were observed with compounds lacking electron donors at the ρ-position of the aryl ring, specifically those possessing the tetralin core.

4.5 Biological Assay Screens

In addition to ligands specifically designed using the above rationale to serve as agonist for ERα, compounds that had been previously generated for our AMI library that fit the profile were

evaluated as well. The increase in presence of and accessibility to small molecule chemical libraries for attenuation of physiological process garners the demand for feasible, reliable and time considerate approach to screening. Herein we utilize two screening approaches first chemical complementation as an initial screen and chemiluminescence as a second screen.

4.5.1 Chemical Complementation

Broadly defined chemical complementation is genetic selection highlighted using a small molecule. Though used to describe a variety of experimental systems, the underlying theme of chemical complementation is the marriage of a molecular chemical entity and a "system" wherein genetic selection can be applied to evaluate molecule-system association. Herein we use a three component (yeast three hybrid) system linking the survival of yeast to the ability of a small molecule modulator to activate nuclear receptor transcription. Single and dual component yeast hybrid systems for ligand-receptor survival have been extensively investigated however the three-hybrid system most suits our purpose [2]. Employing yeast as a host is ideal as yeast is an organism void of endogenous nuclear receptors thereby eliminating the concern of alternative competing receptor-ligand interaction. The yeast strain used (Saccharomyces cerevisiae – PJ69-4A) contains Gal 4 response elements (Gal4RE) used to control the expression of genetic selection genes involved in the histidine (HIS3) and adenine (ADE2) biosynthetic pathway. Transformed in the yeast are two fusion proteins, a Gal 4 DNA binding domain (Gal4DBD) fused to a nuclear receptor ligand binding domain (LBD) and a Gal 4 activation domain (GAD) fused to a nuclear receptor co-activator (SRC-1). At the core of chemical complementation binding of a ligand to the nuclear receptor LBD promotes binding of the LBD: Gal4DBD fusion with the Gal4RE thus prompting a conformation change of the receptor allowing for recruitment of the GAD:SRC-1 fusion. Completion of this cascade actuates transcription of selective genes

(HIS3 or ADE2) thereby allowing yeast to survive in media lacking the presence of histidine or adenine.

4.5.2 Chemiluminescence

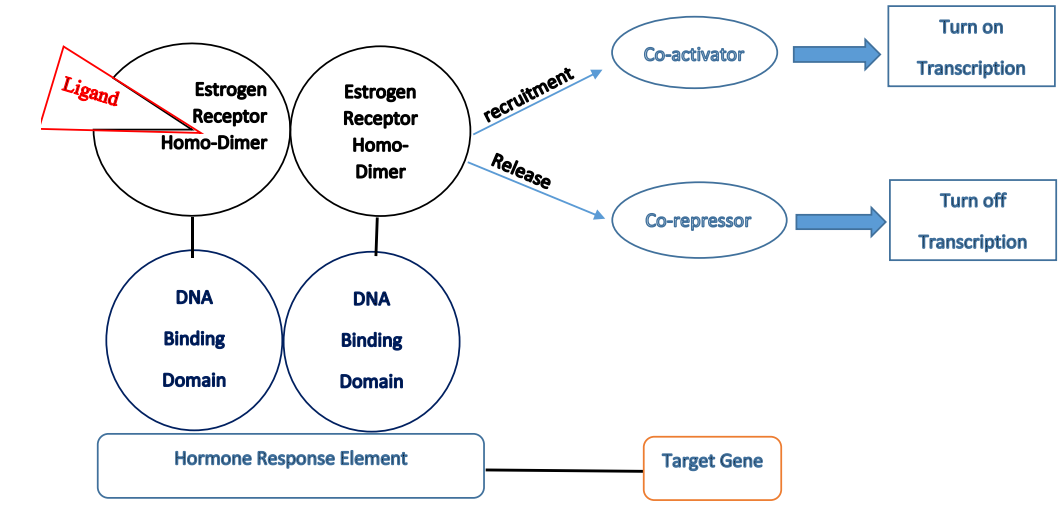
Chemiluminescence is a process by which a chemical reaction generates sufficient energy to produce an excited stated resulting in photon emission where upon relaxation to its corresponding ground state emits light. When the photon emitting reaction is enzymatically driven, the process is specifically designated bioluminescence with the enzyme termed luciferase and the photon emitting substrate termed the luciferin. Largely used as a genetic reporter expression of luciferase can be affixed to a cellular event linking bioluminescence to a quantitatively observable process detectable by a signal. Cellular based assay's formatted to exploit bioluminescence are widely accepted as a viable methodology of high throughput screening compound evaluation. Herein we utilize a luciferase reporter gene assay to measure the efficacy of our ligands with estrogen receptors α and β as well as additional nuclear receptors.

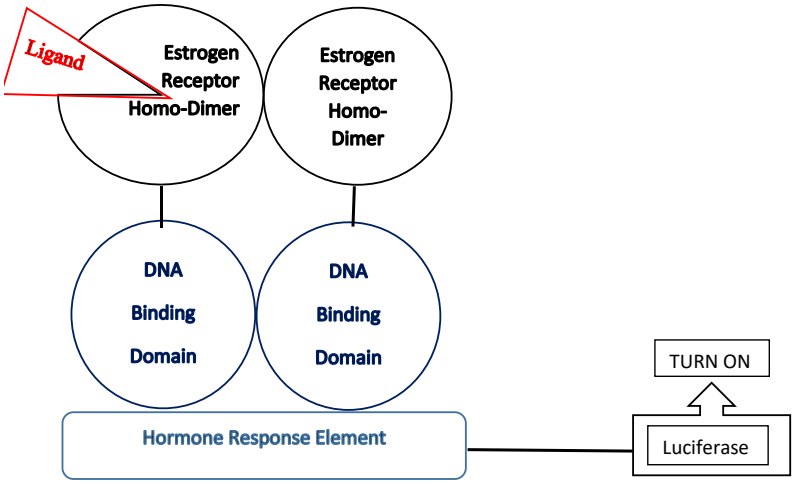
4.5.3 Nuclear Receptors and Reporter Gene Assay

Luciferase reporter assays have been employed with success for receptors such as G-protein coupled receptor and nuclear receptors. In the case of nuclear receptors cloning of the receptors regulatory region of interest takes place upstream of the luciferase reporter while downstream of the reporter houses a promoter with a hormone response element or Gal4 binding domain and a mammalian promoter. In turn ligand binding of the nuclear receptor in the proximity of the promoter activates luciferase gene expression. Firefly and Renilla luciferase are commonly used for this form of assay with the latter being used as a co-reporter. Firefly luciferase functions combining beetle luciferin with ATP, to form luciferyl-AMP as an enzyme bound intermediate.

This intermediate reacts with O_2 to create another bound intermediate, oxyluciferin, in a high-energy state. For our assay purpose wild-type estrogen receptor α ligand binding domain was fused to Gal4DBD, this plasmid was then cloned into mammalian expression vector pCMX. This plasmid fusion operated under the control of a cytomegalovirus (CMV) promoter. The reporter plasmid, p17*4TATAluc, contained the Renilla luciferase gene under the control of four Gal4 response elements located upstream from a minimal thymidine kinase promoter served as co-plasmid for dual reporting. The pCMXβgal, a plasmid containing the β-galactosidase gene under the control of the mammalian CMV promoter, was also used as an internal standard.

A)





B)

C)

Figure 4.9 Screening analysis method for evaluating GFP inspired agonist with nuclear receptors. A) Chemical complementation via the application of the yeast three hybrid system. Yeast survival is a direct result of the ligands ability to bind and activate the nuclear receptor transcription machinery. B) Luciferase assay production of the enzyme luciferase is controlled by the ligands ability to bind to the nuclear receptor. Luciferin in the presence of luciferase generates light C) Reaction that characterizes the production of light in the luciferase assay. Firefly luciferin acts as the substrate broken down by the enzyme luciferase to produce light.

4.6 Generation I Results

4.6.1 Chemical Complementation

AMI ligands based on rational drug design as described above were evaluated for activation of estrogen receptor α using chemical complementation as an initial screen. With the employment of the Gal 4 protein, a ligand independent transcription factor, two fusion proteins were constructed: Estrogen Receptor α ligand binding domain (ERαLBD) was fused to the Gal 4 DNA binding domain (DBD) generating the ErαLBD:Gal4DBD protein fusion. Estradiol (E2) the endogenous ligand for the estrogen receptor was tested and exhibited growth resulting in an optical density (OD) of 0.5 ± 0.04. To ensure observed responses was restricted to the ligand, the vehicle (ethanol:dmso) designated no ligand or NL was tested in which no growth was shown yielding an OD of 0.04 ± 0.02. The Gal4 protein (intact) was used as a positive control providing an OD of 0.6 ± 0.1. Of ligands designed to investigate hydrogen bonding ability only activated growth was observed for CW32 producing an optical density (OD) of 0.8 ± .1.

Screening ligands centralized on exploration of ligand volume did not result in observed activation, however interesting comparison can be made from CW32 and CW30. Observed activation can be noted in CW32 where the only contrast is conversion of the isopropyl to a propyl group supporting the implication positive influence from isopropyl groups.

Compounds designated to evaluate molecular topography also produced no new led indicating that lowering the overall polarity of the compound for this type of ligand may not directly result in enhanced activity. Sensitivity for both AMI's and estradiol were measured via dose response curve with ligand concentration ranging from 10^{-5} to 10^{-12} M. Estradiol is several orders of

magnitude more sensitive than CW32. Much of the subset did not induce growth however the discovery of activity with CW32 prompted further screening.

Table 4.4 Chemical Complementation Results of 1st Generation Ligands. Results of 1st generation compounds indicating growth in ligand CW32 and control estradiol (E2). No growth was observed in the vehicle consistent growth was observed in the positive control Gal 4.

Ligand	**Optical Density**	**Fold Induction**	**Gal 4**
Vehicle	0.04 ± .001	1	0.6
E2	0.5 ± 0.03	11	0.6
CW32	0.8 ± 0.1	21	0.6

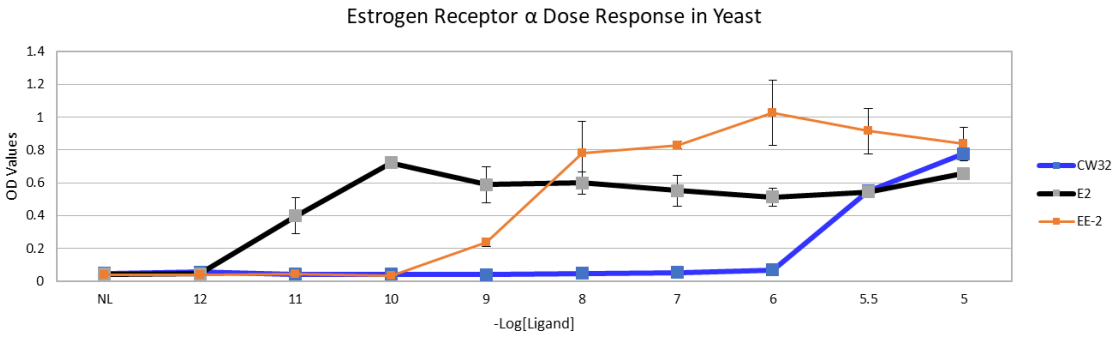


Figure 4.10 Generation I dose response curve in yeast. Active compound CW32 (blue) and controls estradiol (E2 black) and ethynyl estradiol (EE-2 orange) were included. CW32 is comparable in potency however not in sensitivity.

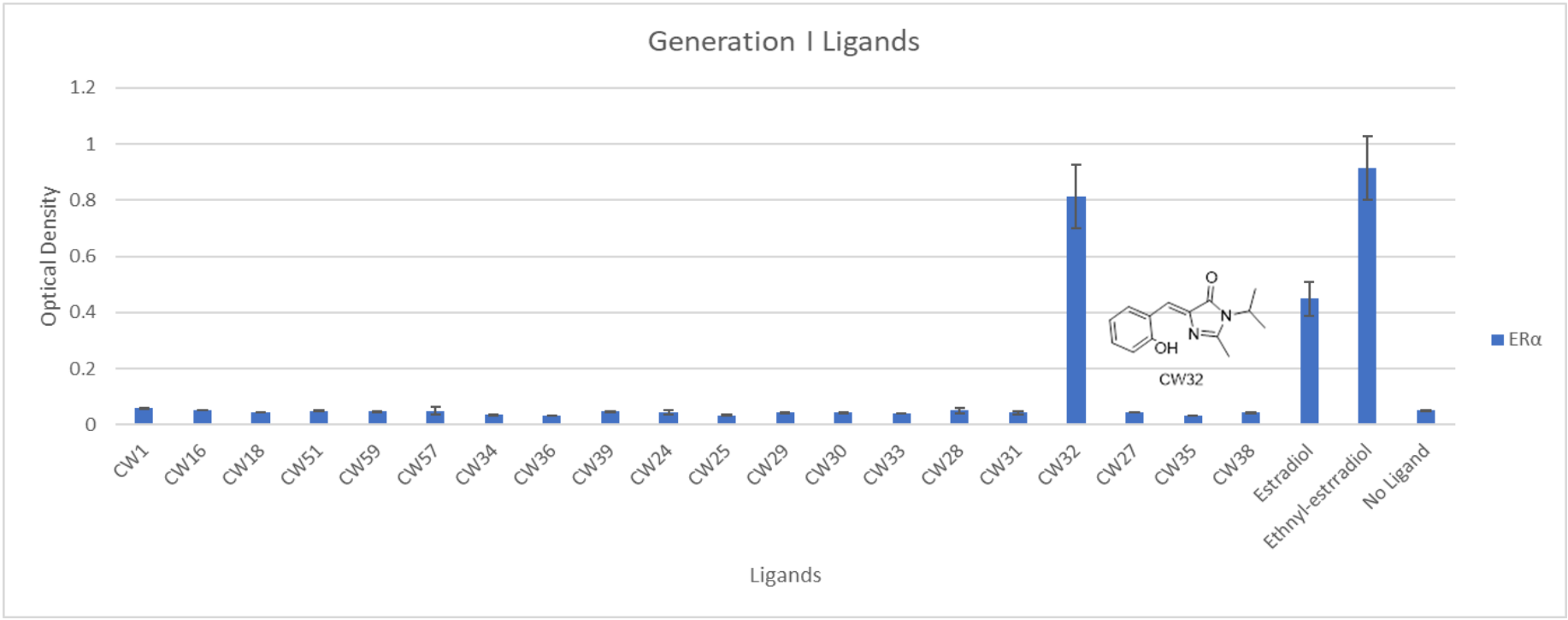


Figure 4.11 Chemical complementation screening of generation I assay. Initial screening in yeast hybrid assay resulted in display of activation in CW32.

4.6.2 Luciferase Results

To determine whether observed activation in yeast existed within mammalian cell, human embryonic kidney 293T (HEK293T) cells were transfected with the plasmid pCMXwithERαLBD. Like the fusion protein used in chemical complementation this plasmid contains the Gal4DBD (GBD) fused to the wild-type ERα ligand binding domain LBD. The pCMXβgal, a plasmid containing the β-galactosidase gene under the control of the mammalian CMV promoter, was also used as an internal standard. Ligands were added, and cells were harvested 48hrs after the addition of ligand and analyzed for luciferase and β-galactosidase activity. All data points represent the average of triplicate experiments normalized against β-galactosidase activity, standard deviations are represented using error bars.

Utilizing the bioluminescence assay the initial subset of AMI ligands screened in yeast were assayed in human embryonic kidney 293T cells (HEK293T) for comparison purpose. Ligand responses are reported as fold inductions (quotient of ligand response over vehicle). Estradiol (E2) was added as a control and to ensure observed responses were restricted to the ligand, the vehicle (ethanol:dmso) was tested (NL = no ligand) and normalized to 1. E2 exhibited fold induction of 26 ± 0.7. Ligands exemplifying responses 90% or greater than that of the E2 were deemed high activators, 75-89% as moderate and 74% or below as modest.

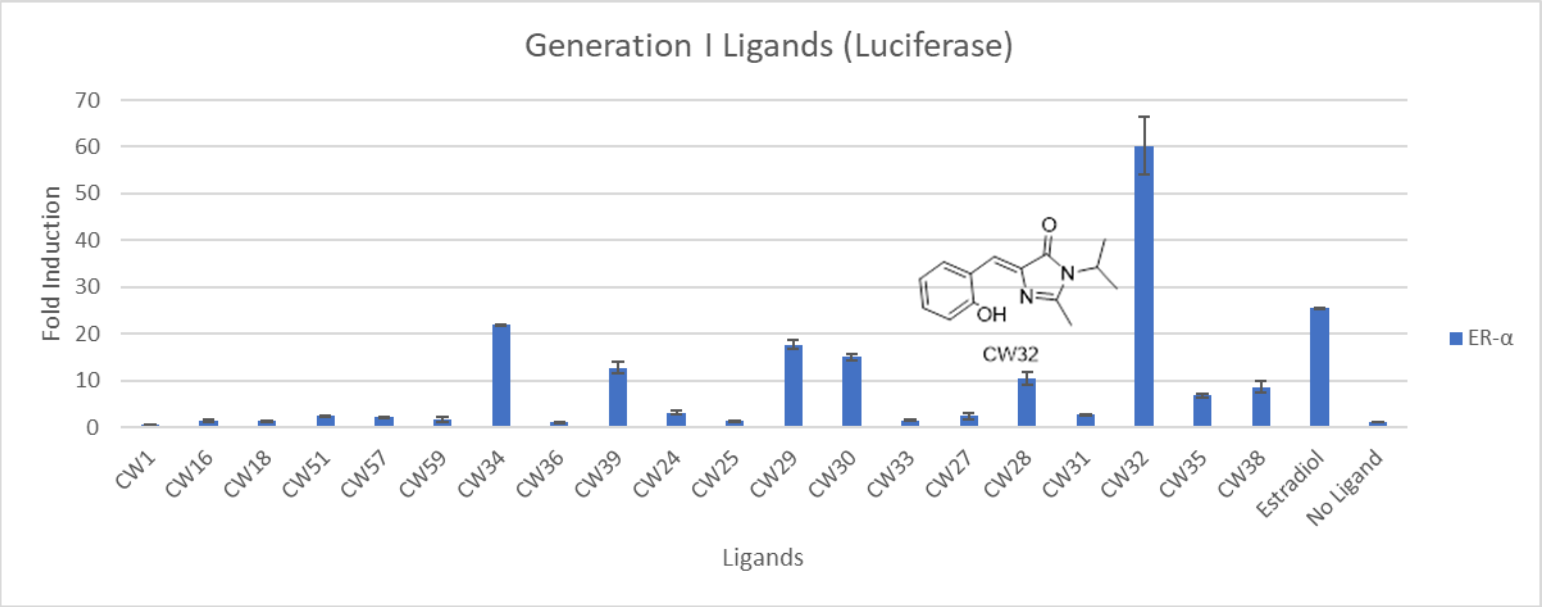


Figure 4.12 Luciferase assay generation I ligands screen. Results indicated 6 ligands that demonstrated activity for ERα.

4.6.2.1 Hydrogen Bonding Network

Recall the hydrogen bonding network can be viewed as a function of three principal amino acids Gln-353 and Arg-394 at the rear of the ligand binding pocket and His-524 at the front of the pocket. Ligands that were generated to explore preservation of a hydrogen bonding network revealed activation from several compounds with a range of potency from modest to high. Ligands CW24 – CW33 were designed to explore interactions with Gln-353 and Arg-394. Compounds CW25 and CW27 both of which possess phenols however due to steric crowding the hydrogen is inaccessible for donation resulted in fold activation of 1.49 ± 0.1 and 2.4 ± 0.6 respectively. CW33 which completely removes the core (A-Ring) phenol abolishes all activity producing a fold induction of 1.5 ± 0.1. This data supported our belief that the core hydroxyl plays a physical donor-acceptor role and is not reduced to solely electronic contribution. Compounds CW28, CW29, CW30, CW31, CW32 all of which possess phenols as accessible hydrogen bond donors, all displayed activation except CW31 which produced a fold induction of 2.6 ± 0.1. CW28 display the lowest fold induction of the group at 11 ± 1.4. Both ligands CW29 and CW30 are comparable in fold induction 18 ± 1.1 and 15 ± 0.7 respectively. Ligand CW32 produced the highest induction of the group at 60 ± 6.2 more than double the response of E2. Although ligands CW28 and CW32 only differ in substitution pattern of the phenol there is a pronounced distinction in activation levels with CW32 displaying induction value 6-fold higher. CW24 possess a methylated hydroxyl rendering the ability of the core phenol to solely hydrogen bond acceptor yields a fold induction of 3.1 ± 0.3. While this fold activation is slightly higher than that of non-activating ligands it does not compare to that of ligands with donor capabilities. Screening of compounds designated to explore the role of the core (A-ring) phenol with Gln-353 and/or Arg-394 illustrated the necessity for a hydrogen bond donor.

Table 4.5 Exploration of the role of the A-ring hydroxyl. Ligands designed to explore the role of the A-ring hydroxyl as a hydrogen bond donor or receptor reveals an activation is observed when the A-ring hydroxyl moiety acts as a donor rather than acceptor.

Ligand	**Fold Induction**	**Donor/Acceptor Role A-Ring**
CW24	3.1 ± 0.3	Acceptor
CW25	1.4 ± 0.1	Donor (Sterically Hindered)
CW27	2.4 ± 0.6	Donor (Sterically Hindered)
CW28	11 ± 1.4	Donor
CW29	18 ± 1.1	Donor
CW30	15 ± 0.7	Donor
CW31	2.6 ± 0.1	Donor
CW32	60 ± 6.2	Donor

Ligands CW34, CW36, CW39, were screened to assess if interaction with His-524 are necessary and if so are donor or acceptor capabilities optimal. All three ligands have a phenol on the A-ring in the para position and an aryl phenol on the nitrogen. CW34 and CW36 are both hydrogen bond donors varying only in the substitution pattern of the N-arylated phenol with the first being para and the latter ortho. CW34 produced a fold induction of 22 ± 0.1 while moving the phenol to the ortho position (CW36) annulled all activity with a fold induction of 1.1 ± 0.1. Lastly CW39 is identical to that of CW34 except the noncore phenol is capped restricting ability to only hydrogen bond acceptance. Removing hydrogen bond acceptor capabilities significantly diminished fold induction to 13 ± 1.3. It appears better activity is achieved with a hydrogen bond donor however distance limitations are applicable.

Table 4.6 Hydrogen Donor vs Acceptor Role of Ligand with His-524. Fold induction as a function of substituents designed to assess the ligands role as a hydrogen donor or acceptor as it relates to histidine-524. Activation is achieved with both the ligand as a donor or acceptor however enhanced activation is seen with the donor.

Ligand	**Fold Induction**	**Role with His-524 Donor/ Acceptor**
CW34	22 ± 0.1	Donor
CW36	1.1 ± 0.1	Donor
CW39	12 ± 1.3	Acceptor

4.6.2.2 Molecular Topography

Attenuation in molecular topography was achieved by incorporation of tetralin cores CW01, CW07, CW16, CW18, CW46, CW47 or napthyl cores CW51, CW57, CW59, CW60. As reported in Table no activation was observed in ligands bearing tetralin or napthyl additions. It is important to note that modification in the scaffold can also contribute to variations within ligand volume and size restrictions.

Table 4.7 Perturbation of Ligand Core. Variation effect on polar surface area Against Activation for different ligand cores. ACD cores promote the best activation of the scope of cores we considered for our ligands.

Ligand Core	**Polar Surface Area**	**Log P**	**Fold Induction**
I. Tetralin			
CW 7	35	4.6	1
CW 18	55	4.7	1
II. Napthyl			
CW51	35	3.3	2.4
CW57	55	3.2	2.1
III. ACD/AMI			
CW32	55	1.9	60
CW34	75	1.6	22
IV. ABCD			
E2	40	3.4 (3.3*)	26

4.6.2.3 Ligand Volume

Ligand volume and size prove to be an interesting dynamic mainly because variation to explore hydrogen bonding network or molecular topography often yields alteration in ligand volume-shape. Several factors have been observed within ligand volume perturbations, table illustrates the volume and consequently effect on fold induction for active ligands thus far.

Table 4.8 AMI Molecular Volume & Fold Induction. The above table demonstrates the relationship between molecular volume and fold induction displaying as molecular volume increases past that of estradiol a decrease in activation can be seen. ***denotes experimental value**

Ligand	**Molecular Volume**	**Fold Induction**
Estradiol (E2)	268 (245)*	26 ± 0.7
CW32	230	60 ± 6.2
CW34	259	22 ± 0.1
CW39	277	12 ± 1.3
CW63	268	13 ± 1.7
CW38	270	8.6 ± 1.2
CW35	253	6.8 ± 0.4
CW33	222	1.5 ± 0.1

In general, it's clear that the closer in volume to estradiol an increase in potency can be observed, however after estradiol's volume a significant drop-off in activation occurs. In ligand families were volume is conserved as is the case of isomers for instance CW32 and CW30 which differ in N-alkylation isopropyl vs propyl respectively yields distinct differences in activation. The same is observed with isomeric family CW32, CW31, CW28, were volume remains unchanged however reposition of the hydroxyl moiety on the aromatic ring can demonstrate superior, abolished or diminished fold induction as demonstrated in table. Ligands CW35 and CW38 explored if the potency enhancement observed from spacing of the isopropyl substituent could be the start of a larger ring system. Ligands CW35 and CW38 exhibited modest activity with fold induction values of 6.8 ± 0.4 and 8.6 ± 1.2 respectively the lowest response of the subset. Indicating that the isopropyl spacing has a limitation and is not further enhanced by generating a larger cyclic system.

Table 4.9 Isomeric AMI Molecular & Fold Induction. The above table illustrates the effect of rotating the phenol in the A-ring on isomers of CW32 removing interaction with the receptor results in abolishment of activity.

Ligand	**Molecular Volume**	**Activation**	**Receptor H-Bonding**
CW32	230	60.2 ± 6.2	Leu-346
CW28	230	10.5± 1.42	Arg-394, Glu-353
CW31	230	2.64 ±0.12	None

4.7 Structure Activity Relationship Compounds

Ancillary AMI ligands were synthesized by exploiting key activation criteria provided from 1st generation results. Activating compounds were examined against compounds of similar nature that did not activate to devise a structure activity relationships diagram. The AMI ligand can be divided into three portions for examining and iteration development: A-ring, core, and D-ring.

4.7.1. A-Ring

The hydroxy on the A-Ring proved to be not only a necessity for activation but also selective in its position and relationship. Comparison of compounds CW33 against CW29, CW30, or AB18 were the only structural distinction is the presence of a hydroxyl group illustrates the need for the moiety to induce activation. The relationship for the hydroxyl as a hydrogen bond donor is clearly defined by the lack of activation in compounds that possesses the hydroxyl but limit its role to hydrogen bond acceptor (CW24). Lastly within isomeric families such as CW28, CW31, CW32 and AB18, CW30, CW29 shift in activation can be viewed as a direct consequence of preference in hydroxyl position.

4.7.2. Core-Scaffold

Modifications in the core structure demonstrated the complexity with this class of ligand for generating compounds adhering to volume restrictions while minimizing polarity. Comparison of unsubstituted tetralin, naphthyl, aryl methylene and indole (previously generated) cores revealed aryl methylene imidazolone core to have the lowest molecular volume. This is beneficial because it allows for more substitutions to be made while still heeding to pocket volume restrictions this

aided in our decision to focus predominately on the aryl methylene imidazolone core. While focus was given to aryl methylene core compounds without inclusion of other parameters such as hydrogen bond contribution of the A-ring the possibility that additional types of cores including indoles and naphthyl cannot be discarded if properly attenuated.

4.7.3. D-Ring

Activation appears to be tolerant to several D-Ring substituents with preference for small nonpolar groups such as propyl and isopropyl, small cyclic rings and hydroxylated aromatic rings. Comparing D-Ring substituents of compounds all bearing a 4-hydroxyl substitution on the A-Ring provides the effect from the substituent. Compounds CW28, AB18, CW35, CW34, CW38, CW39. CW34 show aromatic hydroxyl rings contribute with the largest activation of the set however with a significant contribution coming from aromaticity and hydroxyl as compared with CW38. The second highest activation of the above set was generated by CW39 demonstrating that while preference may be for a hydrogen bond donor incorporation of a hydrogen bond acceptor still produces activation higher than that of compounds lacking either capability. Additional substituents isopropyl (CW28), propyl (AB18), cyclopentane (CW35) all having comparable activation.

Consideration was still giving to our 3 established criteria of hydrogen bonding, topology and ligand volume while constructing derivatives.

4.7.4 Hydrogen Bonding Network

To maximize hydrogen bonding network, additional hydroxyl moieties were incorporated in the A-ring phenol generating a series of diols. Increasing the number of hydroxyl group on the core phenol should ideally increase the number of hydrogen donor-acceptor contacts between critical

amino acids and the ligand thus strengthen the hydrogen bonding network ultimately resulting in enhanced binding and activation.

4.7.5 Ligand Volume and Size

Effort to increase activation while maintaining general ligand volume and size was conducted by combining active D-Ring substituents with active A-ring cores. Using cores substitution that 1st generation compounds have shown to be preferred along with front end substituents that have shown to be favorable should result in optimal compounds. As illustrated in the figure below with the core of CW32 and the substituents of CW34 to generate CW61.

4.7.6 Molecular Topology

Due to the incorporation of additional hydroxyl moieties and hydroxylated aromatic rings on existing polar cores the molecular topology (polar surface area) will increase. The premise is that the modifications being made are more favorable and create interactions that neglect the effect that increasing polar surface area may have.

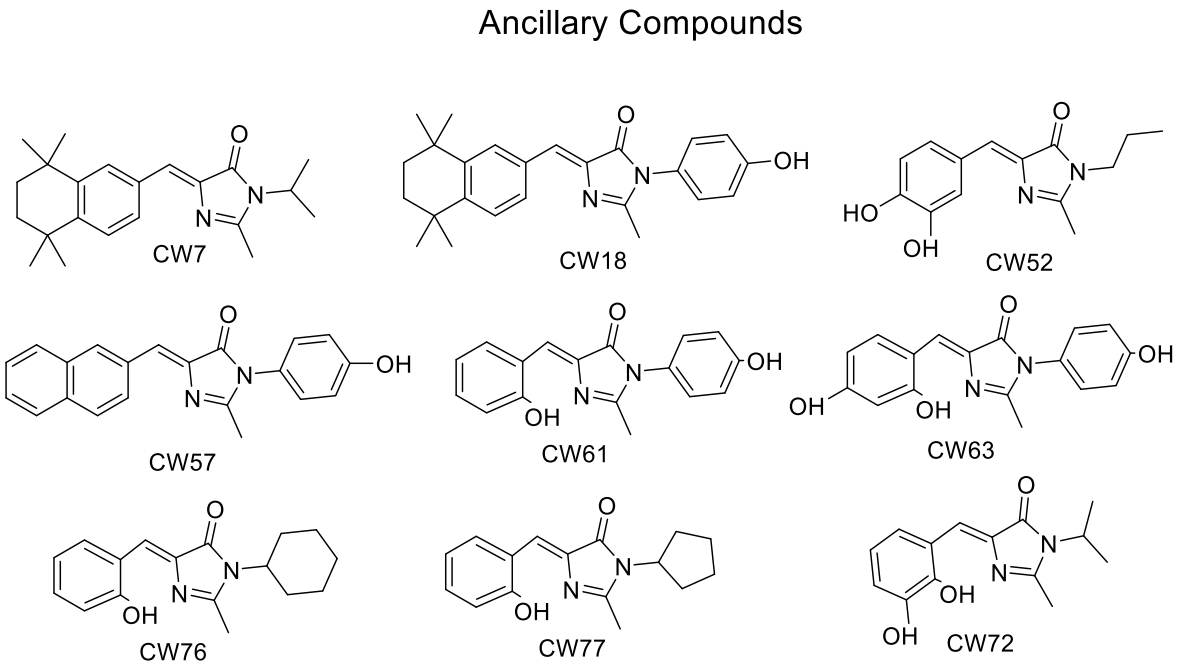


Figure 4.13 Strutrue activity relationship compounds. A 2nd generation of compounds were synthesized after data anaylsis of exploration compounds.

4.8 Generation II Results

In accordance with the biological screening assay protocol explained in section 4.5 of this chapter 2nd generation compounds were tested initially via chemical complementation and secondly via luciferase assay.

4.8.1 Chemical Complementation

AMI ligands based on 1st generation structure activity relationship studies were evaluated for activation of estrogen receptor α using chemical complementation as an initial screen. As detailed previously estradiol (E2) the endogenous ligand was used as a positive control and exhibited an optical density (OD) of 0.5 ± 0.1. Precaution was taken to ensure that observed responses was restricted to the ligand, the vehicle (ethanol:dmso) designated no ligand or NL was tested in which no growth was observed yielding an OD of 0.04 ± 0.04. The Gal4 protein

(intact) was used as an additional positive control given an OD of 0.7 ± 0.1. Chemical complementation screening of 2nd generation resulted in no observation of activation from any of the next generation compounds.

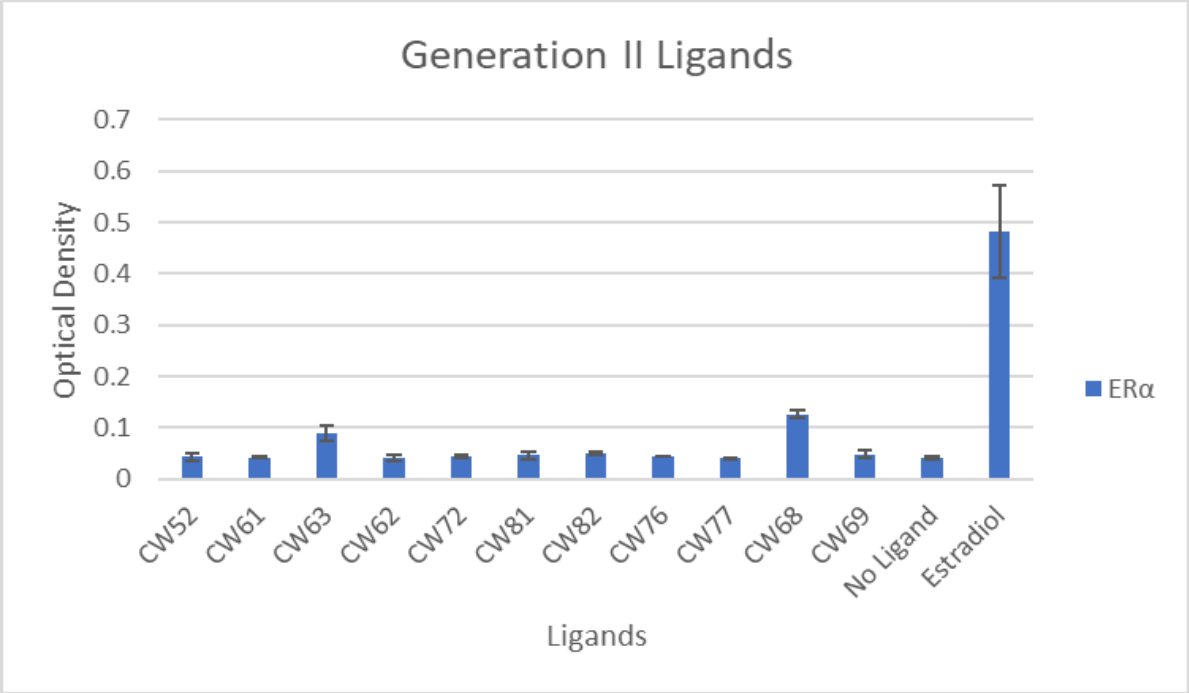


Figure 4.14 2nd Generation chemical complementation results. Evaluation of ligands designed from generation I results.

4.8.2 Luciferase Assay

Utilizing chemiluminescence as a screening tool previously outline in this chapter 2nd generation compounds were screened via luciferase assay. It is important to note that just because activation was not observed within the chemical complementation yeast assay that does not guarantee no activation will be observed in this assay as well.

4.8.2.1. Hydrogen Bonding Network

Multi-substituted 3,4-diol (CW52) and 2,3-diol (CW72) were synthesized with n-propyl and isopropyl alkylations respectively 2,4-diol (CW63) an analog of CW34 was also synthesized.

Interestingly the construction of the diols resulted in active compounds for CW63 and CW72 but lead to significant loss in activity from their parent compounds with fold induction values of 13 ± 1.7 and 5.4 ± 0.4 whereas the combination of hydroxyl to the para and meta position generating CW52 abolished all activity. This leads us to believe that the incorporation of more diols does not ensure strengthening of the hydrogen bonding network and in turn more potent compounds.

4.8.2.2. Ligand Volume -Size

In an initiative to further understand the role of volume with activating compounds specifically the effect of small nonpolar substituents such as the isopropyl in CW32 a series of analogues were synthesized. These analogues all shared the ortho phenolic ring as in CW32 however varied in amide substitutions but only containing substituents of active compounds. This series generated CW61, CW76, and CW77 containing p-aminophenol, cyclohexyl and cyclopentyl substituents respectively. While activity remained in CW61 the shift in the core phenol caused fold induction of 13 ± 3.3, dramatic decrease in activation from CW34 and CW32. All activity was abolished with ligands CW76 and CW77 confirming a preference for small non-polar substituents.

4.8.2.3. Molecular Topology

Lastly to gather insight regarding molecular polarity as it pertains to the top two activators CW32 and CW34. The tetralin analogues CW7 and CW18 were synthesized as well as napthyl analogue CW57 were synthesized. Based upon previous results it was not surprising that the napthyl nor tetralin cores produced no response.

Dose responses were conducted on top active compounds to determine the degree of sensitivity in comparison to that of estradiol. Ligand concentration ranged from 10^{-5} to 10^{-12} M.

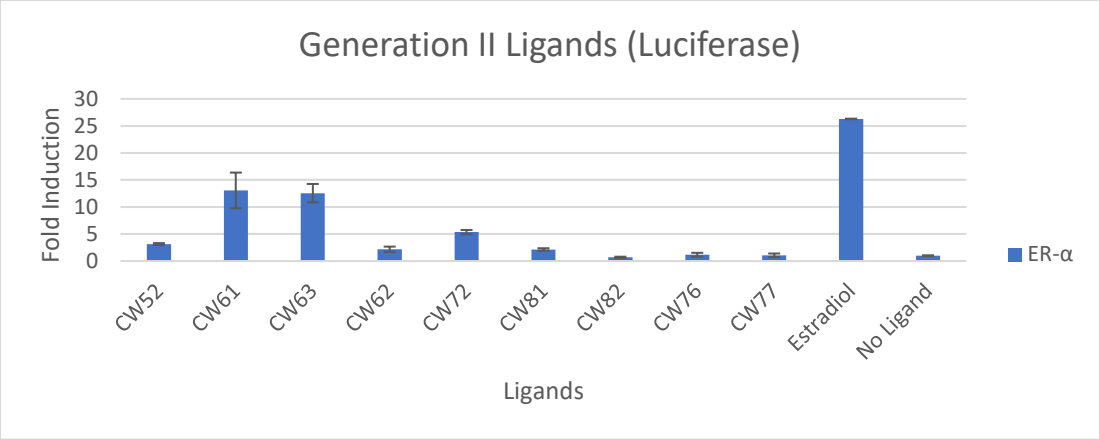


Figure 4.15 2nd generation ligands screened in luciferase assay. Ligands designed from the development of SAR of 1st generation were tested for activity.

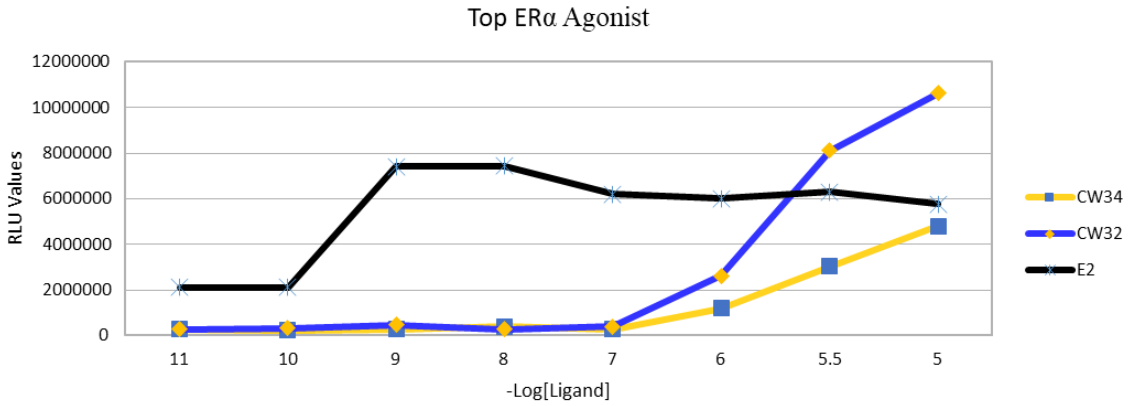
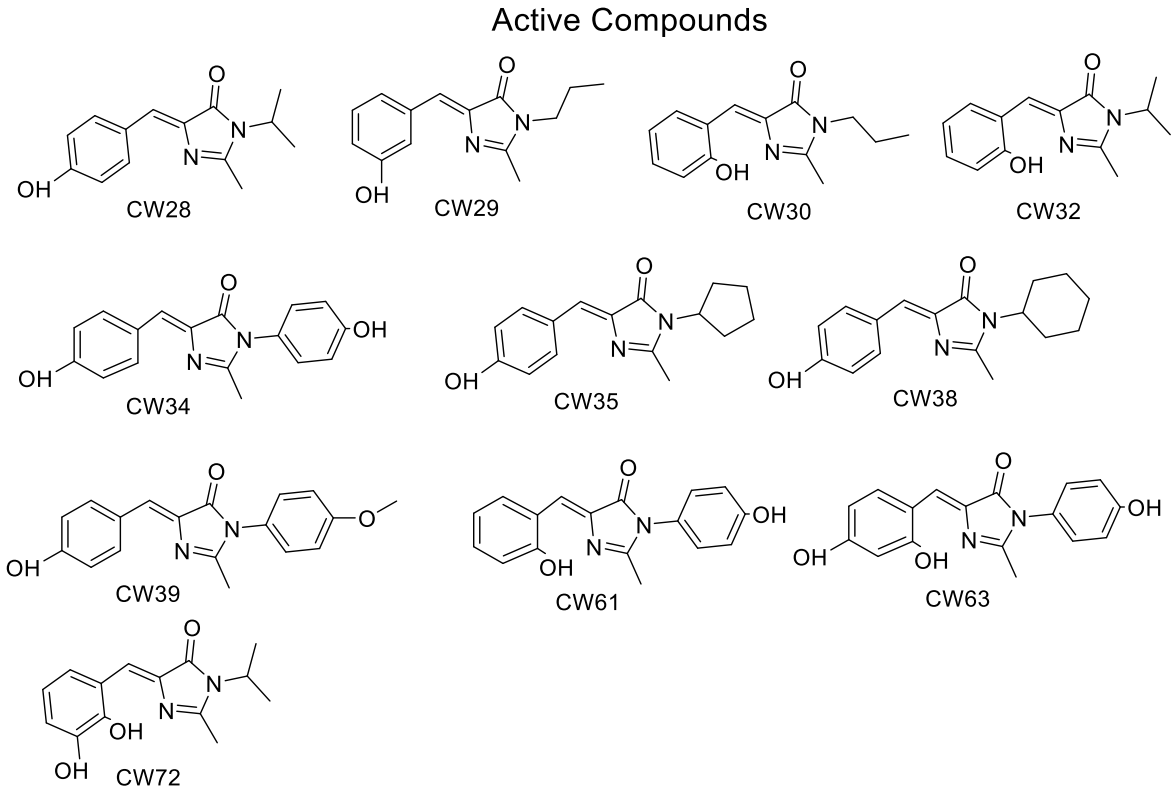


Figure 4.16 Dose Response for top ERα agonist. Generation I and II screening in luciferase assay revealed CW32 (blue) and CW34 (yellow) as top activators. While high in potency estradiol remains more sensitive.

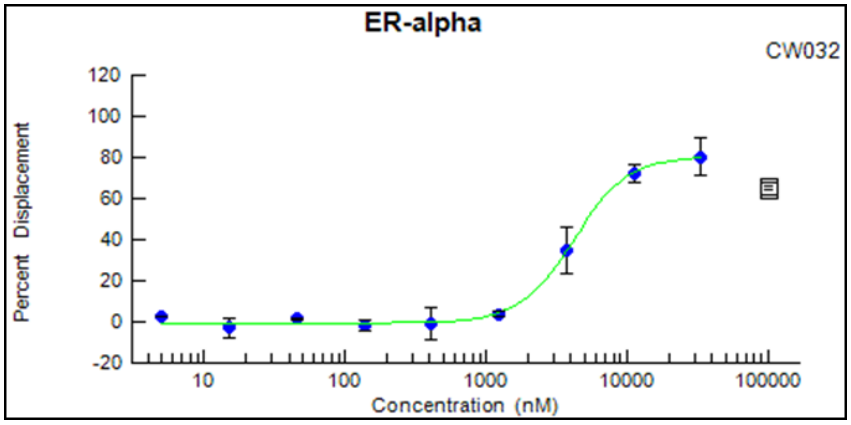
To correlate the activity in vivo with in vitro and binding analysis, the top two activators (CW32, CW34), a modest activator (CW72) and a non-activator (CW33) were subject to testing using TR-FRET (time resolved fluorescence resonance energy transfer) competitive binding assay. This assay employs estrogen receptor α labeled via a GST tag bound to its anti-GST antibody that is terbium labeled. Competitive binding to estrogen receptor α occurs via our AMI agonist ability to displace a fluorescent ligand (tracer) in our case labeled estradiol (ES2 Green tracer).

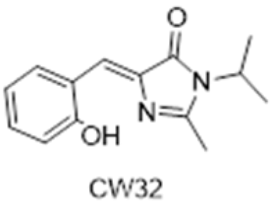
The proximity of labeled ligand and the tagged antibody directly corresponds to the strength of the FRET signal. Here successful displacement of the tracer by our AMI agonist will result in a decrease in FRET signal. Results of a 10-point titration produced dose response curves that are within agreement with our observations. The highest activator CW32 caused the largest displacement at the highest concentration, followed by CW34 and only modest displacement observed with CW72. Lastly CW33 demonstrated no displacement which was expected based upon HEK chemiluminescence results.



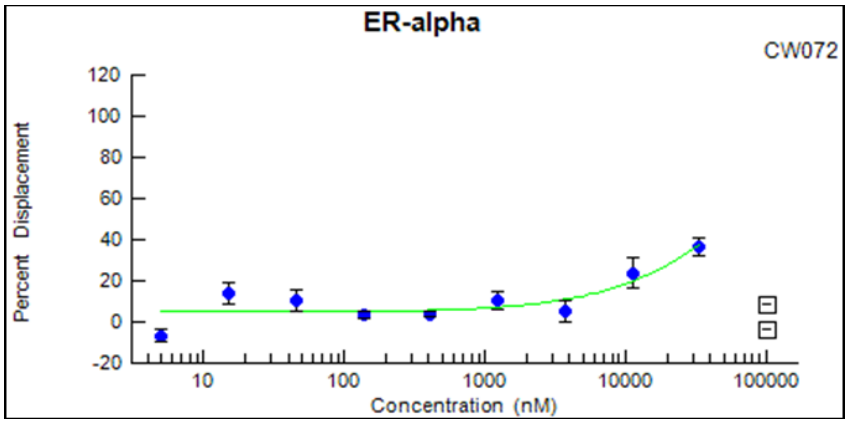
Ligand	Fold Induction	Response Percentage
Vehicle (Ethanol : DMSO)	1.00 ± 0.007	NA
CW28	10.5 ±1.42	41.1
CW29	17.7 ± 1.06	69.2
CW30	15.1 ± 0.702	59.0
CW32	60.2 ± 6.2	235
CW34	21.9 ± 0.091	85.6
CW35	6.82 ± 0.424	26.6
CW38	8.58 ± 1.22	33.5
CW39	12.7 ± 1.32	49.8
CW61	13.1 ± 3.3	51.1
CW63	12.6 ± 1.7	49.1
CW72	5.36 ± 0.39	20.9
Estradiol (E2)	25.6 ± 0.732	100

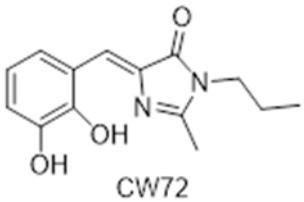
Figure 4.17 Structure of GFP chromophore inspired active ligands and fold induction. Active structures (above), table shows fold induction and response as percentage to estradiol.





CW32





CW72

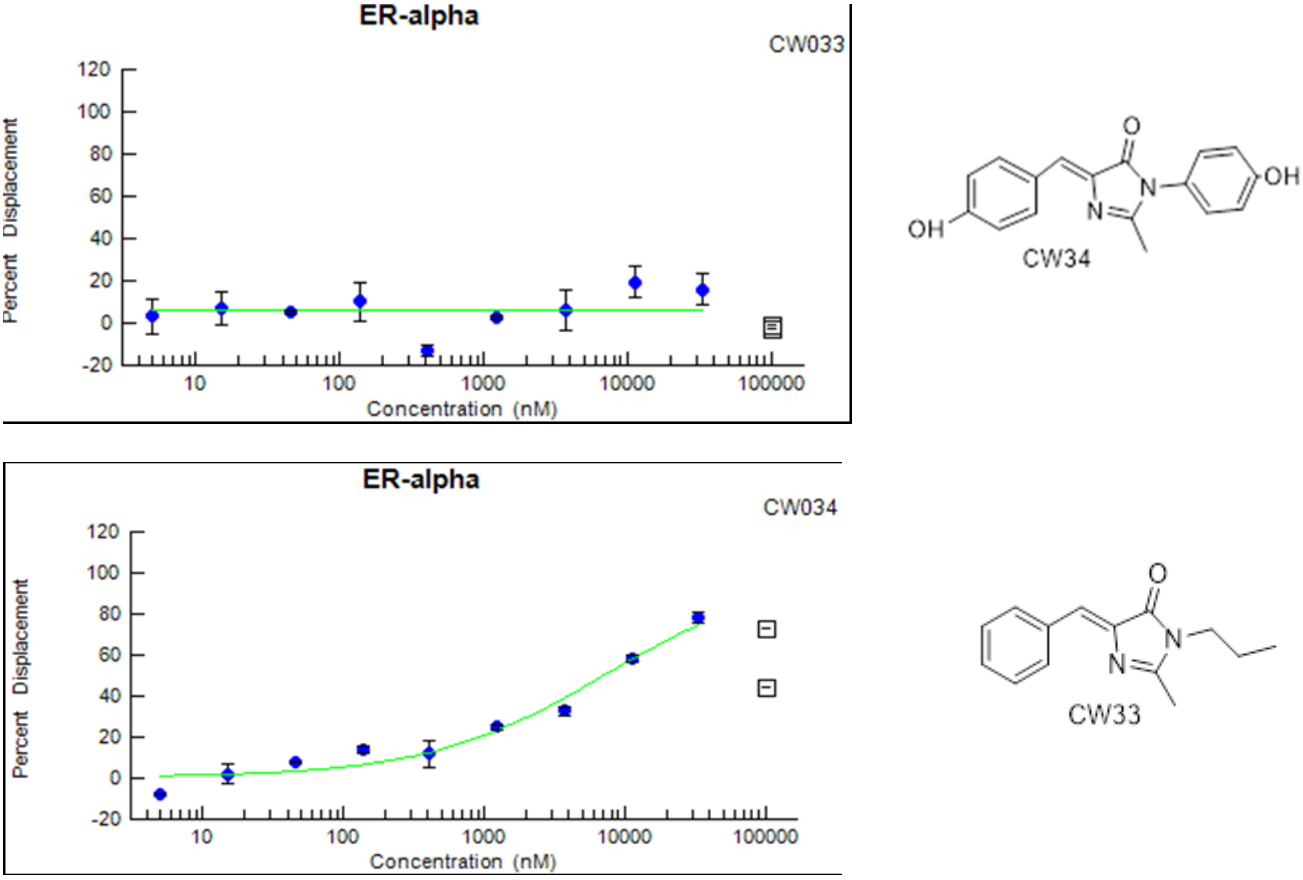


Figure 4.18 TRFRET results of top activators and lowest activator. TRFRET results of top activator and lowest activator confirmed luciferase assay results of binding with ligands successfully displacing estradiol. CW33 was used as a negative control (non-binder) which showed no displacement.

4.9 Ligand Activation Rationale

As previously shown in literature ligands that stimulate a response from the estrogen receptor either as agonist or antagonist arise from a diverse background of compounds however all share commonalities highlighted as essential for activation. In our exploration of GFP chromophore inspired compounds as a viable class of estrogen receptor agonist main guidelines supported in literature were observed to identify the role played with this class of molecule. In designing this class of ligand attention was giving to the following 3 criteria: 1) preservation of hydrogen bonding network via key amino acid residues within the ligand binding pocket 2) size and

volume occupied by the ligand and lastly 3) molecular landscape of the ligand with respect to polarity.

4.9.1 Hydrogen Bonding Network

The modulation of the hydrogen bonding network between AMI ligands and the estrogen receptor α ligand binding pocket demonstrates a pronounced effect on activation. While possible weak polar interactions with the aromatic ring contribute to activation, the ability of the phenol to participate in hydrogen bond donation seems to be priority for activation. Docking of several active AMI ligands indicate ligand-residue contacts that preserve the role of the phenol moiety with known residues Arg-394 and Gln-353. The phenol in ligand CW28 exhibits hydrogen bond donation reminiscent to estradiol's A-ring, where it is located $2.6A^3$ and $3.7A^3$ from Arg-394 and Gln-353 respectively. The lack of this hydrogen bond network resulted in a lack of activity in compounds CW25 and CW27 where the phenol is inaccessible due to steric hindrance and CW33 where the phenol is non-existent. This led to the rational that preserving the hydrogen bond network and ensuring that the hydrogen bond network is sterically favorable, results in activation. Ligand CW24 bears a methylated phenol revoking hydrogen bond donation ability and consequently abolishing activity thus supporting the role and need of a hydrogen bond donor on the A-ring. While some of the active ligands do not maintain the Arg-394 or Gln-353 contacts the H-bonding donation network remains consistent. Case in point are modulators CW32 and CW30 were the phenol hydrogen participates in H-bonding with the backbone carbonyl of Leu-346 solely in CW32 and cooperatively with Gln-353 in CW30. It can be concluded that to achieve activation with this class of ligand a hydrogen bonding network between the ligand and ligand bonding pocket must be established. Our data supports that this

network can be inclusive of amino acids residues or their backbones. Not only is this relationship a requirement but our evidence supports it is the most important contributing factor.

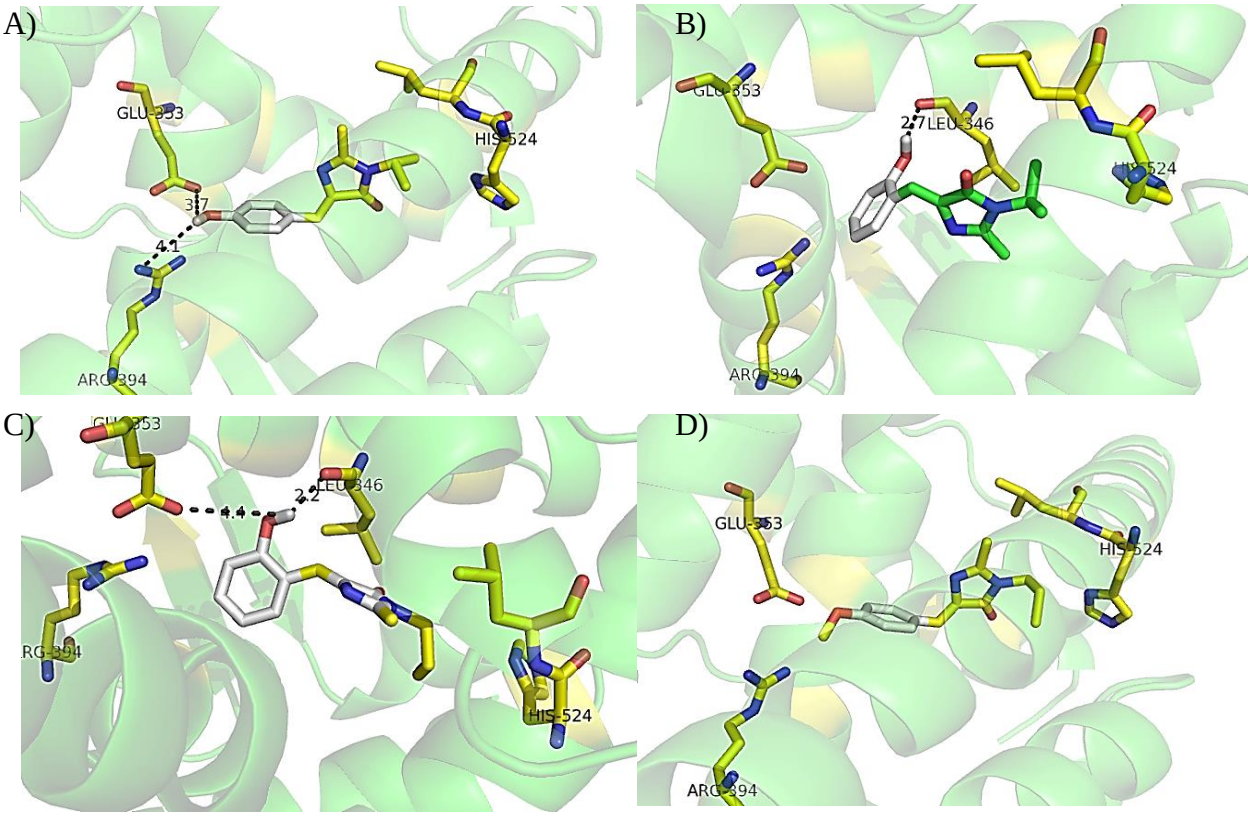


Figure 4.19 Modeling of the hydrogen bonding network between AMI ligands and the ERα:LBD. **A**) Phenolic A-Ring in CW28 participates in H-bonding with ARG-394 and GLU-353 residues. **B**) Phenolic A-Ring in CW32 participates in H-bonding with LEU-346 backbone carbonyl. **C**) Phenolic A-Ring in CW30 participates in joint H-bonding with GLU-353 and carbonyl of LEU-346. **D**) In ability of phenolic A-ring within CW24 to participate in H-bonding.

The hydrogen bonding network is extended to the front of the AMI ligand involving His-524. Linked to the 17β position of estradiol, His-524 ability to serve as a hydrogen bond donor has been highlighted as a key residue within the LBP [37]. The imidazolone carbonyl in our ligands possess the same ability as the 17β secondary alcohol in estradiol to participate in hydrogen bond accepting. Comparative distance of the carbonyl to His-524 residue yields results farther than that of estradiol relationship however it is possible that this process is water mediated. The carbonyl distance of CW28, CW30 and CW32 are $4.8A^3$, $5.5A^3$ and $5.1A^3$ respectively. It is

worth noting that the carbonyl (AMI) and alcohol (E2) are not in perfect alignment, likewise the carbonyl is rotationally restricted as compared to a secondary alcohol. The role of the oxidation of a secondary alcohol to a carbonyl is most notably demonstrated in the conversion of estradiol to estrone although this conversion resulted in diminished activity from that of estradiol [38]. It can be postulated in the case of these ligands that the carbonyl has a positive contribution with the resonance structure of the amide leading to a more enhanced hydrogen bond acceptor thus keeping the hydrogen bond network intact.

In addition to external hydrogen bonding between the ligand and the LBP a second type of hydrogen bonding may exist solely involving the ligand. Intramolecular hydrogen bonding can be present amongst certain ligands in this class case in point CW32. The nitrogen incorporated in the imidazolone ring and the phenol in ortho-substituted AMI ligands are positioned to participate in intramolecular hydrogen bonding. This potential preservation of hydrogen bonding network is more evocative of estradiol's B-ring and may provide prospective into discrepancies in isomeric activation. Reduction in activation occurs within CW32 series as the phenolic hydroxyl group shifts from ortho to para consequently as the intramolecular bonding distance between the hydroxyl group and the nitrogen increases. Analogue CW31 where the hydroxyl group is in the meta position and devoid of the ability to hydrogen bond with itself or the LBP all activity is annulled. The resurgence of activity with CW28 is a testament to the prominence of the need for hydrogen bonding network (external) however the contrast in activation between CW32 and CW28 may speak to the contribution of intramolecular hydrogen bonding.

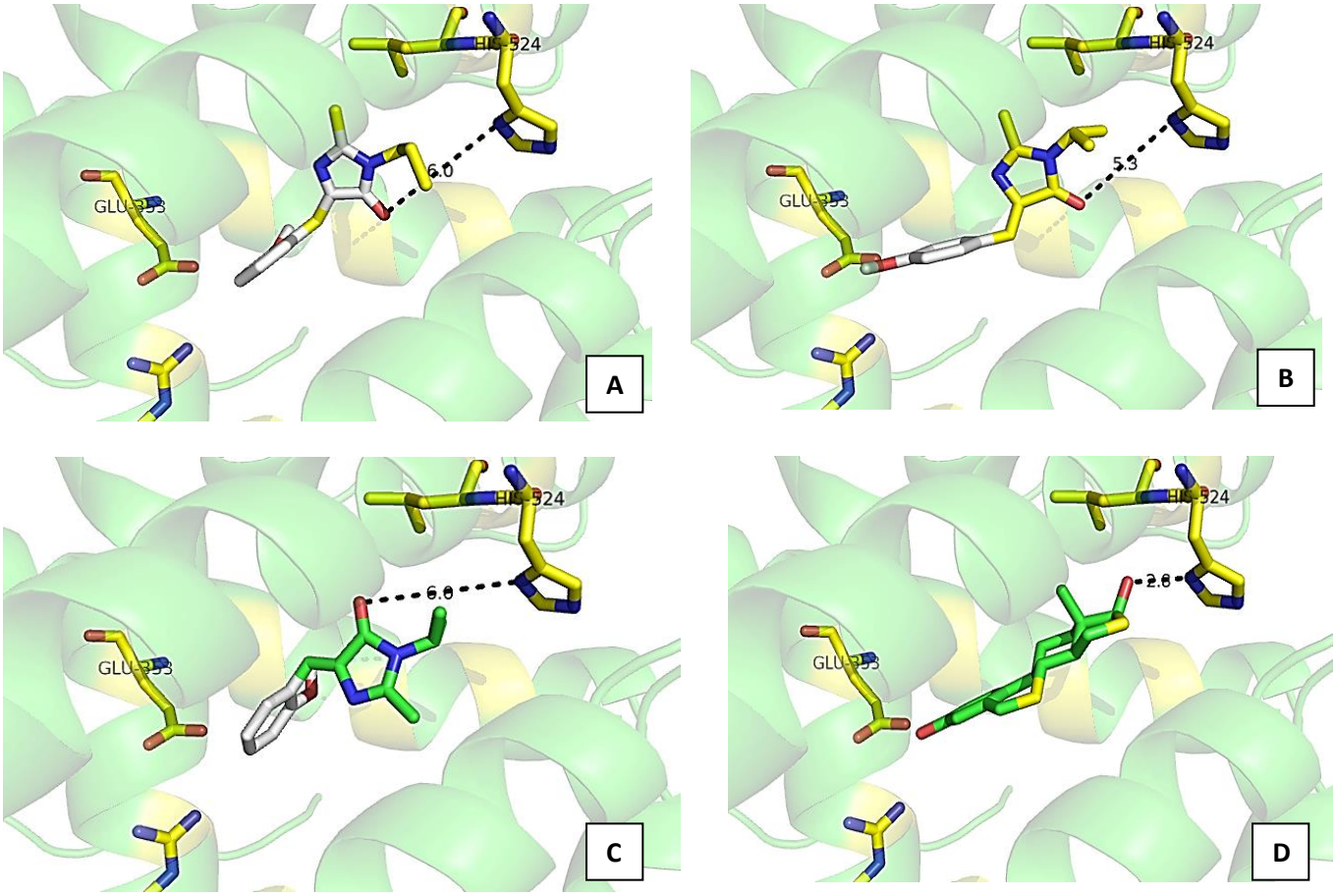


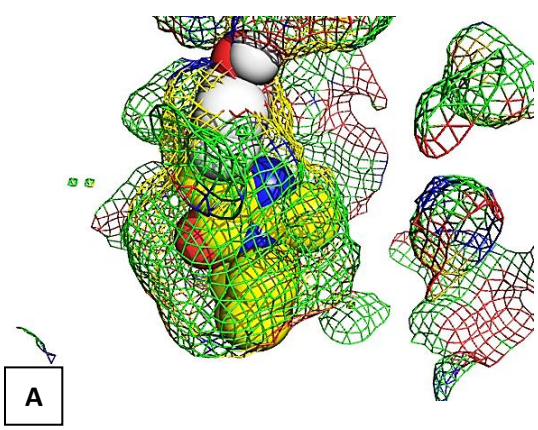
Figure 4.20 Exploration into the role of His-524 with AMI ligands comparative to that of E2.
A) Carbonyl of CW28 interacting possibly via water mediated process with His-524 residue.
B) Carbonyl of CW30 interacting possibly via water mediated process with His-524 residue.
C) Carbonyl of CW32 interacting possibly via water mediated process with His-524 residue.
D) Hydroxyl of E2 preserving H-bonding network with His-524 residue.

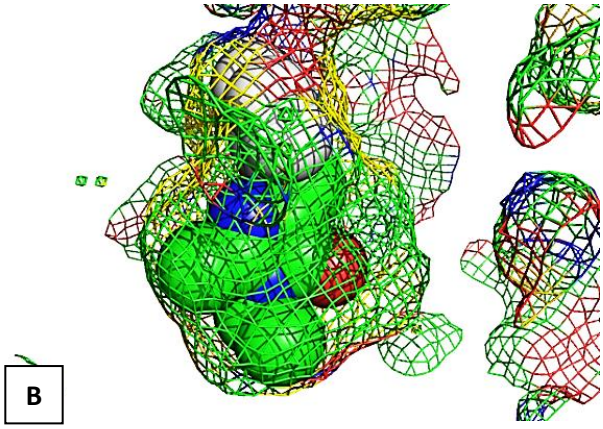
4.9.2 Ligand Volume(Shape/Size)

Molecular volume and estrogen receptor α activation for this class of ligand display an inverse relationship. Our data suggest that as molecular volume increases activity diminishes. Using molinspiration cheminfomatics software [34] molecular volumes of select active and non-active AMI ligands were calculated. Favorable molecular volume for this class of ligand is roughly half the available space of the LBP volume (450A^3). This finding is consistent with the experimental

molecular volume of estradiol ligand at 245A^3. In our data set top activator CW32 is comparable with a predicted value of 230A^3. The role of molecular volume can be viewed in two perspectives within a series of isomeric compounds and across a family of compounds. Within an isomeric family such as CW28, CW29, CW30, CW32 where volumes are identical at 230A^3 hydrogen bonding network preservation becomes the determining factor of activation.

However, when looking across families of compounds for instance CW32, CW34, CW39 molecular volume play a significant role following an inverse relationship. Anomalies within this trend are observed across families wherein a ligand larger in volume maybe more active than a smaller one as shown in the case of CW35 and CW63, which includes. Rationale for this can be provided by discrepancies in hydrogen bonding ability. Interestingly compounds that have ideal molecular volume CW33 at 222A^3 but lack hydrogen bonding displays no activity. It can be concluded for this class of ligand that while hydrogen bonding ability is essential for activation to occur, molecular volume can be viewed as an enhancing factor that possess threshold requirements that diminishes activation if violated.





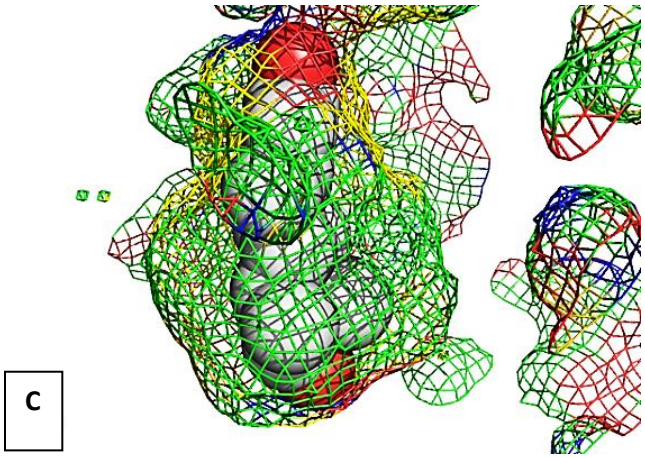


Figure 4.21 Modeling of volume and size of AMI ligands in comparison with that of E2. A) Illustrates the volume of the pocket (mesh) occupied by E2. B) Illustrates the volume of pocket occupied by CW32. C) Illustrates ligand CW35 perturbing out of the pocket.

4.9.3 Molecular Topology-ACD ring effectiveness

Topology also plays a substantial role in binding with the LBP being largely hydrophobic and size restricted. ERα crystallography complexes with ligands, and site mutagenesis studies indicate a hydrophobic LBP intolerant to polar substitutions [35]. Using Katzenellenbogen pharmacophore as our guideline certain substituents had to be included locking in a specific base level of polar surface area. The stipulation of a phenolic A ring is consistent with what we observed for all active AMI ligands. The incorporation of hetero-cyclic rings to provoke estrogen receptor modulation is not unprecedented and has been employed in studies to show modest relative binding affinities [36]. It is plausible that the modest affinity is a direct result of the introduction of a higher polar topography. As expected the inclusion of the imidazolone in our C ring scaffold adds polarity elevating our ligands polar surface area to that above estradiol, but whether the deviations are considerable or minuscule are dependent upon the amide alkylation's and aryl substitutions. The D ring as expressed by Katzenellenbogen pharmacophore as generally aromatic is consistent with several of our active compounds with higher levels of fold activation observed for aromatic rings CW34, CW39, CW63 and CW61 then for non-aromatic rings CW35 and CW38. Consequently, the addition of the aromatic ring increases the polar surface area. While CW32 does not contain a secondary aromatic ring but an isopropyl as a direct result the

polar surface area is lowered although still above that of estradiol. The subset did contain analogues that were lower in polar surface area than estradiol for example CW7 at 35A however deviations from pharmacophore features and a high molecular volume prevent activation. Currently we can conclude that for this class of ligand the small polar aromatic cores are best when in conjunction with small non-polar substituents, however at present an optimal molecular topology has not been determined.

In addition to polar surface area the effectiveness of the ACD ring system plays an integral role. The figure below illustrates that within the pocket the ACD ring system shares alignment commonalities with that of the indigenous ligand estradiol. The hydroxyl moiety of the GFP inspired agonist such as CW32 aligns with that of the methyl of estradiol and faces in the same plane as the hydroxyl in estradiol. Good alignment occurs within the A-ring of both ligands even though the hydroxyl moieties do not align. This strengthen the rationale behind the usage of the ACD ring system and its interactions within the estrogen receptor.

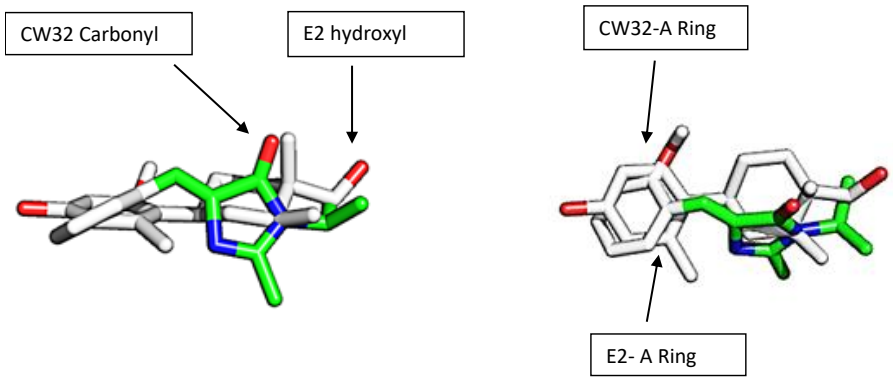


Figure 4.22 Alignment of CW32 and estradiol in binding conformation in the ligand binding domain.

Additional support for the GFP inspired chromophore as estrogen receptor agonist claim can be found by examining the distance between specific atoms. The interatomic distance of the oxygens A-ring phenol and secondary phenol in estradiol has been shown to be at an optimal spacing to promote residue interaction within the ligand binding pocket. The arylmethyleneimidazolone core agonist adopts similar atom spacing with the A-ring and the D-ring substituents. As highlighted in table 4.10 certain D-ring substitutions can achieve this interatomic distance allowing the ACD ring system to access the same interactions as that of the indigenous ligand. As highlighted in the table below and illustrated in figure 4.22 CW34 is one such ligand that the interatomic distance from the phenol (A-ring) to that of the phenol (D-ring) is like that of estradiol. Similar observations can be made for ligands other ligands including top activator CW32. However, in the absence of a second phenol the carbonyl group on the arylmethyleneimidazolone core can be used as a second marker.

Table 4.10 Interatomic distance and fold induction. Inter atom distance for the GFP-inspired chromophore ligand oxygen to oxygen distance is like that of estradiol. Distance is conserved in non-binders however other factors such as hydrogen bonding play a critical role in activation. * denotes experimental value

Ligand	Fold Induction	Molecule Distance Length (A)
Estradiol	26 ±0.7	10.8 (10.9) *
CW32	60 ± 6.2	7.6
CW34	22 ± 0.1	13.5
CW33	Not active	7.2

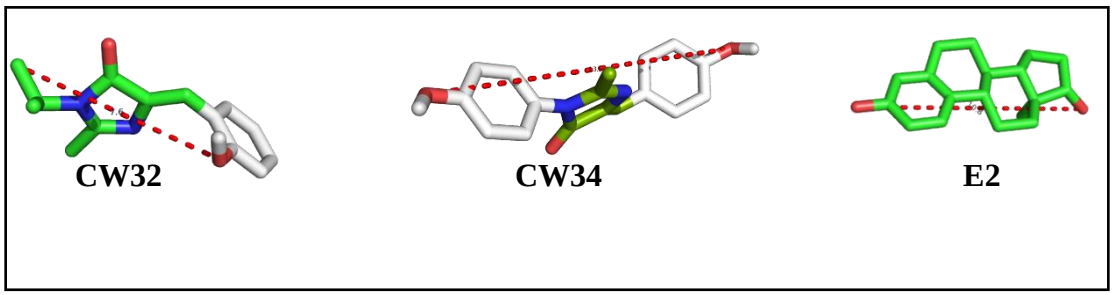


Figure 4.23 Oxygen to oxygen distance for estradiol and top two activators. The interatomic distance is illustrated in red with ligands bearing similar distance but different ring systems.

4.10 SUMMARY

The development of ligands for estrogen receptor α continue to be of high importance as both agonist and antagonist. Herein we have reported the synthesis, evaluation and modeling of a series of GFP-chromophore inspired arylmethylene-imidazolone as estrogen receptor α ligands. Through perturbation of analogues we have been able to determine key criteria for activation of this class of ligand to be hydrogen bonding, molecular volume, and molecular topology. Establishing a hydrogen bonding relationship between the ligand and the ligand binding pocket emerged as the most dominant contributing factor. Without this relationship activation for this class of ligand cannot be achieved. Molecular volume and topography are two additional factors that fluctuate as attenuation to ligands occur. Molecular volume demonstrated a direct relationship with activity we demonstrated for this class of ligand as volume exceeds that of estradiol activity diminishes. Through molecular topography we demonstrated that the most active of this class of ligand contain polar aromatic cores and small nonpolar substituents. Several ligands were found to be active and in agreement with literature described pharmacophore, of specific interest CW32 elicited potency 2-fold above that of estradiol. The interesting potency of CW32 has prompted binding affinity studies as our next investigation. While not as sensitive as estradiol the ability to promote higher activation, levels have intriguing potential contraceptive implications like ethynyl estradiol. In agreement with specific aim 2 focus will now turn on ensuring ligands can promote activation selectively. Testing in ERβ as well as other nuclear receptor families is the next step.

4.11 Materials and Methods

EXPERIMENTAL SECTION

Chemistry

General

Estradiol was purchased commercially from Sigma-Aldrich (St. Louis, MO). All AMI ligands were synthesized using Baldridge's adaptation of Bazureau protocol [24]. Proton nuclear magnetic resonance (1H-NMR) and carbon 13 nuclear magnetic resonance (13C-NMR) were recorded on a Varian Mercury 300 nuclear magnetic resonance spectrometer in appropriate deuterated solvent (CDCl3 or DMSO), chemical shifts are reported in □ units with tetramethylsilane (TMS) as an internal standard. Electron spray ionization (ESI) analyses were performed on and time of flight (TOF) methodology. Thin layer chromatography (TLC) was conducted on pre-coated silica gel plates. Chromatography was conducted using silica gel unless otherwise stated. Spectroscopic data as well as yields and melting points can be found in supplementary data. Solvents were used as is unless stated as dry then which solvent was distilled over appropriate drying agent and store under nitrogen over molecular sieves.

General Preparation of Schiff Base (I). To an equivalent of aldehyde was added amine in slight excess (1.1 equivalent) with stirring overnight. Reactions were conducted as neat only when both reagents were liquid, in instances where either the aldehyde or amine were solid 2mL of ethanol was used as a solvent. In most cases after 12-14hrs the product was afforded as a white or colored precipitate that was collected via filtration and washed with small amounts of cold ethanol. Alternatively, several reactions afforded oils that were concentrated in vacuo resulting in nearly quantitative yields.

General Synthesis of Imidate (II). In 125mL of ether was suspended potassium carbonate (5.5g, 39.8mmol) and glycine methyl ester hydrochloride (4.99g, 39.8mmol). To the suspension was added 20mL of water followed by ethyl acetimidate hydrochloride (4.92g, 39.8mmol). The mixture was shaking vigorously with venting for 6 minutes followed by decanting of the organic layer. An additional 100mL of ether was added and shaken continued further for 6 minutes. The organic layers were combined and washed with brine, dried over anhydrous magnesium sulfate; and filtered. The filtrate was concentrated in vacuo to give a clear oil. The imidate was used without further purification. The imidate was stored under nitrogen at 0^{o}C for no more than a week.

General Synthesis of Substituted Arylmethyleneimidazolonone (III). To an equivalent of Schiff base in 2mL of ethanol was added the previously synthesized imidate (1.02 equivalent) and stirred overnight. Precipitation occurred overnight for most reactions and afforded a solid that was filtered, washed with cold methanol and dried in vacuo. In the absence of a precipitate, the crude reaction was placed at 0^{o}C for 48hrs upon which crystallization occurred. When crystallization did not afford product, chromatography was conducted on the crude reaction mixture yielding pure product as a solid or oil.

Preparation of 2,5-dichloro-2,5-dimethyl hexane (IV) 28. To 240mL of reagent grade concentrated hydrochloric acid was added 2,5-dimethyl-2,5-hexanediol (15.6g, 106.9mmol) and stirred overnight at room temperature. After stirring 500mL of water and 500mL of methylene chloride was added to the crude mixture. The organic layer was separated, and the aqueous phase extracted with methylene chloride. The organic layers were combined washed with water; and dried with anhydrous magnesium sulfate; and filtered. The filtrate was concentrated in vacuo to give a white solid (18.0g, 92%).

Preparation of 1,1,4,4,6-pentamethyl-1,2,3,4-tetrahydronaphthalene (V) 29. A stirred solution of 2,5-dichloro-2,5-dimethyl hexane (2.50g, 13.65mmol) in dry toluene (50mL) was cautiously treated with aluminum trichloride (180mg, 1.35mmol) and refluxed under dry conditions for 12 hours. After which the crude solution was cooled to room temperature and poured over ice. The organic layer was separated and washed with 10% HCl, water, brine, dried with anhydrous sodium sulfate and filtered. The filtrate was distilled to remove excess toluene and collected as an oil (2.45g, 89%). ^{1}H NMR δ 1.30 (s, 12H, 4CH_3),1.69 (s, 4H, CH_2CH_2) 2.35 (s, 3H, CH_3) 6.99 (d, 1H Ar-H) 7.19 (s, 1H, Ar-H) 7.31 (d, 1H, Ar-H). ^{13}C NMR δ 21.6, 32.3, 32.4, 34.3, 34.5, 35.6, 35.7, 126.8, 126.9, 127.3, 135.0, 142.1, 144.9.

Preparation of 5,5,8,8-tetramethyl-5,6,7,8-tetrahydronaphthalene-2-carbaldehyde (VI) 30. To a solution of V (4.08g, 20.1 mmol) in 20mL of acetic acid heated to 100^{o}C under vigorous stirring was added was added ceric ammonium nitrate (47.54g, 86.7mmol) in 50% acetic acid. Dropwise addition was carried out over 2 hours followed by stirring for 2 hours. The crude solution was cooled and poured over ice. The aqueous layer was extracted with petroleum ether, washed with water, brine, dried with anhydrous magnesium sulfate; filtered and concentrated in vacuo. The crude mixture was purified via column chromatography using 10% ether/hexane as a mobile phase. Purification produced (3.3 g, 76 %). ^{1}H NMR δ 1.28 **(s,** 12H, 4CH_3), 1.71 (s, 4H, CH_2CH_2), 7.41 (d, 1H, Ar-H), 7.57 (d, 1H, Ar-H), 7.81 (s, 1H, Ar-H), 9.89 (s, 1H, H-C=O). ^{13}C NMR δ 31.5, 31.7, 34.1, 34.7, 34.7, 34.8, 126.6, 127.3, 128.5, 134.3, 145.7, 152.2, 191.9.

Chemical Complementation in Yeast

Conventionally chemical complementation is genetic selection highlighted using a small molecule. Herein we use a three-component system linking the survival of yeast to the ability of a small molecule modulator to activate nuclear receptor transcription. The yeast strain used

(Saccharomyces cerevisiae – PJ69-4A) contains Gal 4 response elements (Gal4RE) used to control the expression of genetic selection genes involved in the histidine (HIS3) and adenine (ADE2) biosynthetic pathway. Transformed in the yeast are two fusion proteins, a Gal 4 DNA binding domain (Gal4DBD) fused to a nuclear receptor ligand binding domain (LBD) and a Gal 4 activation domain (GAD) fused to a nuclear receptor co-activator (SRC-1). At the core of chemical complementation binding of a ligand to the nuclear receptor LBD promotes binding of the LBD: Gal4DBD fusion with the Gal4RE thus prompting a conformation change of the receptor allowing for recruitment of the GAD:SRC-1 fusion. Completion of this cascade actuates transcription of selective genes (HIS3 or ADE2) thereby allowing yeast to survive in media lacking the presence of histidine or adenine. (Fig.3). Yeast trans formant (pGBDhERαLBD) was grown overnight in 2mL of media lacking leucine and tryptophan (SC-LW) at 30°C with shaking at 300rpm. After growth, yeast was pelleted (4°C, 7min, 3700rpm,) media discarded and re-suspended in 1mL of water. A second pellet was formed and re-suspend in 1ml of water before a 45µL aliquot was added to 5mL of water. Media lacking histidine, leucine and tryptophan (SC-HLW with 0.1mM 3AT) was prepared and ligand-media mixes were made via dilutions from pre-made ligand stock solutions (10^{-2} M) with a final concentration of 10µM. The incorporation of 3AT (0.1mM) was necessary to suppress background arising from imidazole glycerol-phosphate dehydratase a key enzyme in the histidine biosynthetic pathway. In the case of dose responses serial dilutions were performed starting with 10 µM, solutions ranged from (10^{-5}M-10^{-12}M). Using a multi-channel pipetor 20µL of cells were added to 80µL of ligand in a 96 well plate. To wells not containing cell: ligand mix, 80µL of water was added. Plates were incubated 48hrs (30°C and 300rpm) with optical density (OD) measurements recorded at 0, 24, 48-hour intervals. Optical density readings were taken at a wavelength of 630 nm.

Mammalian Cell Culture Assay

Human embryonic kidney 293T (HEK293T) cells (ATCC, USA) were transfected with the plasmid pCMXwithERαLBD. This plasmid contains the Gal4DBD (GBD) fused to the wild-type ERα ligand binding domain LBD (GBD: LBD fusion under the control of a cytomegalovirus (CMV) promoter). The reporter plasmid, p17*4TATAluc, contained the Renilla luciferase gene under the control of four Gal4 response elements located upstream from a minimal thymidine kinase promoter. The pCMXβgal, a plasmid containing the β-galactosidase gene under the control of the mammalian CMV promoter, was also used as an internal standard. Lipofectamine 2000 (Invitrogen, USA) was used as the cationic lipid and transfection experimental details are described in Taylor et al. Ligands were added to the wells at various concentrations. Cells were harvested 48hrs after the addition of ligand and analyzed for luciferase and β-galactosidase activity. All data points represent the average of triplicate experiments normalized against β-galactosidase activity, standard deviations are represented using error bars.

In Silico Docking Studies

Modeling of ligand-receptor complexes were carried out using AutoDock Vina® [36]. The crystal structure of human estrogen receptor α (PDB ID: 1ERE) [26] was obtained from the research collaborator for structural bioinformatics (RCSB) protein data bank (PDB). Ligands were created in ChemBioDraw Ultra 13.0 transferred to ChemBio3D Ultra where they were minimized using MMFF84 force field and saved as pdb files. Ligands were modified with Auto Dock Tools by adding Gasteiger charges to set the partial charge of each atom and hydrogens on polar atoms. Ligand- receptor docking was viewed using the Pymol molecular graphics system.

4.12 References

1.Y.Takei, H.Ando, & K.Tsutsui (Eds): Handbook of Hormones Subchapter 94G

2.Wallace, Owen B.; Richardson, Timothy I.; Dodge, Jeffrey A. *Current Topics in Medicinal Chemistry*, **2003**, 3, 1663-1680.

3. Ascenzi, P.; Alessio, B.; Marino, M. *Molecular Aspects of Medicine*, **2006**, 27, 299-402.

4.Giguere, V.; Dufour, C.R.; Eichner, L.J.; Deblois, G.; Cermakian, N. Cold Spring Harbor Symposia on Quantitative Biology, **2011**, 57-61.

5.Chang, C.Y.; McDonnell, D. *Clin. Cancer. Res.* **2012**, 18, 6089-6095.

6. Tamrazi, A.;*Molecular Endocrinology.* **2002,** 16(12), 2706-2719.

7.Sever. Richard, Glass. Christopher K. *Cold Spring Harb Perspect Biol.* **2013**, 5,167-09

8. Shabnam Farzaneh and Afshin Zarghi. *Sci Pharm*. **2016**, 84(3), 409-427

9. Ascenzi, P.; Alessio, B.; Marino, *M. Molecular Aspects of Medicine*, **2006**, 27, 299-402.

10. Brozowski, A.M. et al. *Nature*, **1997**, 389, 753-759.

11. Feng, W. *Science*, **1998**, 280, 1747-1749.

12. Moras, D. *Current Opinion in Cell Biology*, **1998**, 10, 384-391.

13. Jordan, V. *Journal of Medicinal Chemistry*, **2003**, 46, 6, 883-908.

14. Anstead, G; Carlson, K; Katzenellenbogen, *J. Steroids*, **1997**,62, 268-303

15. Losordo, D.W. *Circulation*, **1994**, 89, 1501-1510.

16. Deroo, B.J.; Korach, K.S. *Journal of Clinical Investigation*, **2006**, 116, 561-570

17. Henderson, B.E.; Feigelson, H.S. *Carcinogenesis*, **2000**, 21, 427-433.

18. Katzenellenbogen, J. *Journal of Medicinal Chemistry*, **2011**, 54, 5271-5282

19. Katzenellenbogen, J. *Journal of Medicinal Chemistry*, **2011**, 54, 5271-5282.

20.Richardson, T. etal. *Bioorganic & Medicinal Chemistry Letters*, **2007**. 17, 3570-3574

21.Stauffer, S. etal. *Journal of Medicinal Chemistry.* **2000**, 43, 4934-4947

22. Dong, Jian; Abulwerdi, Fardokht; Baldridge, Anthony; Kowalik, Jansuz; Solntsev,Kyril; Tolbert, Laren. *Journal of American Chemical Society*. **2008**, 130, 14096–14098.

23. Baldridge, Anthony; Samanta, Shampa; Jayaraj, Nithyanandhan; Ramamurthy, V; Tolbert, Laren. *Journal of American Chemical Society*. **2011**, 133, 712–715.

24. Baldridge, Anthony; Solntsev, Kyril; Song, Charles; Tanioka, Tatsuro; Kowalik, Janusz; Hardcastle, Kenneth; Tolbert, Laren. *Chemical Communications.* **2010**, 46, 5686–5688.

25. Naumov, Pance; Kowalik, Janusz; Solntsev, Kyril; Baldridge, Anthony; Moon,Jong-Seok; Kranz,Christine; Tolbert, Laren. *Journal of American Chemical Society.* **2010**, 132, 5845–5857.

26. Conyard, Jamie; Kondo, Minako; Heisler, Ismael; Jones, Garth; Baldridge, Anthony; Tolbert, Laren; Solntsev, Kyril; Meech, Stephen. *Journal of Physical Chemistry B.* **2011**, 115, 1571–1577.

27.Wright, J.; Shadnia, H.; Anderson, J.; Durst, T.; Asim, M.; Salfiti, M.; Choueiri, C.; Pratt, M.A.; Ruddy, S.; Lau, R.; Carlson, K.; Katzenellenbogen, J.; O'Brien, P.; Wan, L. *Journal of Medicinal Chemistry*, **2011**, 54(2), 433-448.

28.Bindra, J. S.; Neyyarapally, A. T.; Gupta, R. C.; Kamboj, V. P.; Anand, N. *J. Med. Chem.* **1975**, 18, 921–925.

29.Neyyarapally, A.T.; Gupta,R.C.;Srivastava,S.C.;Bindra,J.SS.; Groveer, P. K.; Setty, B. S.; Anand, N. Antifertility agents. *Indian J. Chem.* **1973**, 11, 325–329.

30.Asim, M.; El-Safiti, M.; Qian, Y.; Choueiri, C.; Salari, S.; Cheng, J.; Shadnia, H.; Bal, M.; Pratt, M. A. C.; Carlson, K. E.; Katzenellenbogen, J. A.; Wright, J. S.; Durst, T. *Bioorg. Med. Chem. Lett.* **2009**, 19, 1250–1253.

31.Baldridge, Anthony; Kowalik, Jansuz; Tolbert, Laren; *Synthesis.* **2010**, 14, 2424-2436.

32.Caldarelli, Antonio; Minazzi, Paolo; Canonico, Pier Luigi; Genazzani, Armando A.; Giovenzana, Giovanni B. *Journal Enzyme Inhibition and Medicinal Chemistry.* **2013**, 1, 148-152.

33.Min, Jian; Wang, Pengcheng; Srinivasan, Sathish; Nwachukwu, Jerome C.; Guo, Pu; Huang, Minjian; Carlson, Kathryn E.; Katzenellenbogen, John; Nettles, Kendall W.; Zhou, Hai-Bing. *Journal of Medicinal Chemistry.* **2013**, 56, 3346-3366

34.Stauffer, S.R.; Sun,J.;Katzenellenbogen,B.S.;Katzenellenbogen,.*Bioorg. Med. Chem*. **2000**, 8, 1293–1316.

35.Fink, Brain E.; Mortensen, Deborah; Stauffer, Shaun; Katzenellenbogen, John. *Journal Chemistry and Biology*. **1999**, 6, 205-219

36. Stauffer, S. R.; Katzenellenbogen, J. A. *J. Comb. Chem.* **2000**, 2, 318–329.

37.Dawson, M; Hobbs, P; Derdzinski, P; Chao, W; Frenking, G; Loew, G; Jetten, A; Napoli, J; Williams, J; Sani, B; Wille, J; SchiffL L. *J. Med. Chem,* **1989**, 32, No.7, 1504-1517.

38.Dong, J.; Abulwerdi, F.; Baldridge, A.; Kowalik, J.; Solntsev, K.; Tolbert, L.M. *J. AM. CHEM. SOC.* **2008**, 130, 14096–14098

39.Baldridge, A.; Samanta, S.R.; Jayaraj, N.; Ramamurthy, V.; Tolbert, L.M. *J. AM. CHEM. SOC.* **2010**, 132, 1498–1499

40. Baldridge, A.; Feng, S.; Chang, Y.T.; Tolbert, L.M. *ACS Comb. Sci.* **2011**, 13, 214–217.].

CHAPTER 5

INVESTIGATING SELECTIVITY WITHIN ESTROGEN RECEPTOR β AND THE NUCLEAR RECEPTOR FAMILY

The emergence of an additional subtype of the estrogen receptor, ERβ altered the lens through which estrogen regulation had been viewed. Encoded by a different gene than that of its predecessor, ERβ was cloned in 1996 from rat ovary and prostate [1]. ERβ can often be found expressed in identical systems, tissues and cells as that of ERα with variation in the ratio of expression levels [2]. Like other nuclear receptors ERβ possess the ability to mediate a physiological response in reply to the presence of a ligand. The coexistence of both subtypes in identical tissues and the ability to bind similar compounds facilitated the need of ligands that can be selective in receptor subtype, tissue and response (agonist vs antagonist) profile. These ligands are termed selective estrogen receptor modulators or SERM's [3]. This class of compound is characterized by their ability to promote different behavior in varying tissues i.e. antagonist in mammary but agonist in ovary. The generation of ligands that can be classified as SERM's are of extreme importance, the ability to facilitate binding leading to activation or repression in a discriminatory fashion may prove instrumental in minimizing unwanted side effects. In an effort to answer the question is binding and activation observed in chapter 4 restricted to ERα active, ligands were screened against ERβ. Additionally the previously synthesized library of GFP inspired ligands was included in ERβ screening for discovery of potential ERβ exclusive led.

5.1 Estrogen Receptor β

The regions and domains of ERβ remain the same as traditional nuclear receptors as described in Chapter 3. Though the domains have the same function, the role of AF-1 in transcription with ERβ is weaker compared to that of ERα, therefore recruitment of transcriptional machinery is heavily dependent upon AF-2 [4]. Structurally at 530 amino acids ERβ is smaller than ERα while fewer amino acids are present ERβ maintains 44% sequence homology, with higher conservations located within specific regions AF-1 (16%), DBD (96%) and LBD (59%) [5,6].

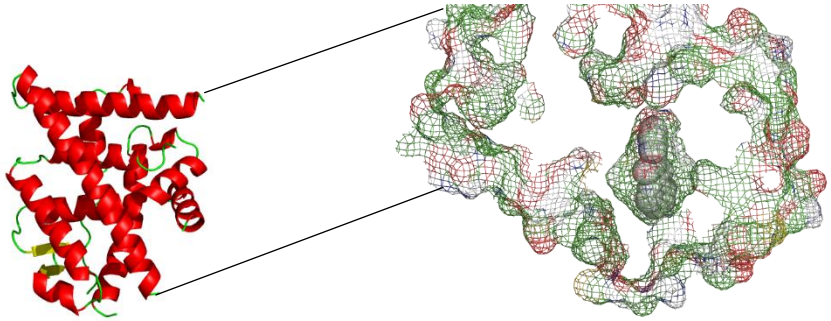
In addition to locations such as cardiovascular system [7], breast tissue and reproductive organs where both subtypes can be located, ERβ can be found exclusively within the colon, lungs, prostate, bladder, gastrointestinal tract and in higher expression levels in ovaries, bones, vascular epithelium and regions of the brain. [5]. The wide spread cellular distribution of ERβ identifies the need for ligand development that is deliberate and careful thus a critical comprehension of targeting the ligand binding pocket is a must.

5.1.1 Ligand Binding Pocket

The ligand binding pocket serves as the processing center for ligand-receptor interactions. The mainly hydrophobic pocket has two defining characteristics the volume of the pocket and the amino acid residues that make up the lining of the pocket. Physically the ERβ ligand binding pocket is $375A^3$ [8] about 100 angstroms smaller than ERα. Lining the available space of the pocket are a series of amino acids residues reminiscent of ERα. Key residues in ERβ ligand binding pocket include His-475, Gln-305, Arg-346 analogous to essential residues His-524, Gln-353 and Arg-394 in ERα. Comparative analysis between the ligand binding pocket of the two subtypes reveals that only a two residue substitutions exist, Leu-384 and Met-421 in ERα are

converted to Met-336 to Ile-373 respectively in ERβ [8]. The presence of similar LBP residues and an established hydrogen bonding network identical to ERα facilitates concern that active ligands may also bind ERβ. The lack of large divergence within the ligand binding pocket makes examining the profile of activating ligands important to begin to discern features that may contribute to selectivity.

Figure 5.1 Estrogen receptor β secondary structure with ligand binding pocket inset. Structure of ERβ predominately helices(red) and loops (green). Inset shows mesh ligand binding domain with



shown ligand binding pocket complexed with genistein.

5.1.2 Ligands

The generic estrogen receptor pharmacophore established by Katzenebollgen [8] holds true for both receptor subtypes therefore ligands such as estradiol activate ERβ in addition to ERα mainly due to the similarity of residues within the pocket therefore similar features will be deemed necessary to promote binding. In addition to binding estradiol with high affinity, [6,9] ERβ can also bind other natural as well as synthetic ligands bound by its counterpart for instance 4-hydroxytamoxifen. Binding of 4-hydroxytamoxifen to ERβ demonstrates pure antagonist behavior where as in certain tissues possessing ERα partial agonist profile is exhibited [7] the variation in behavior is accredited to diversity in pocket size.

Like its predecessor ERβ can bind a diverse series of ligands that share small commonalities such as genistein, liquiritigenin, S-equol and the notable 4-hydroxy tamoxifen. The capability of ERβ to bind ligands like that of ERα, challenges the ability to generate ligands that are selective binders. Most ligands for ERβ contain an ACD ring system analogous to ERα ligands including our GFP inspired chromophore ligands. Though equivalent in ring system diversity in the core is implicated as one of the major pathways to achieving selectivity.

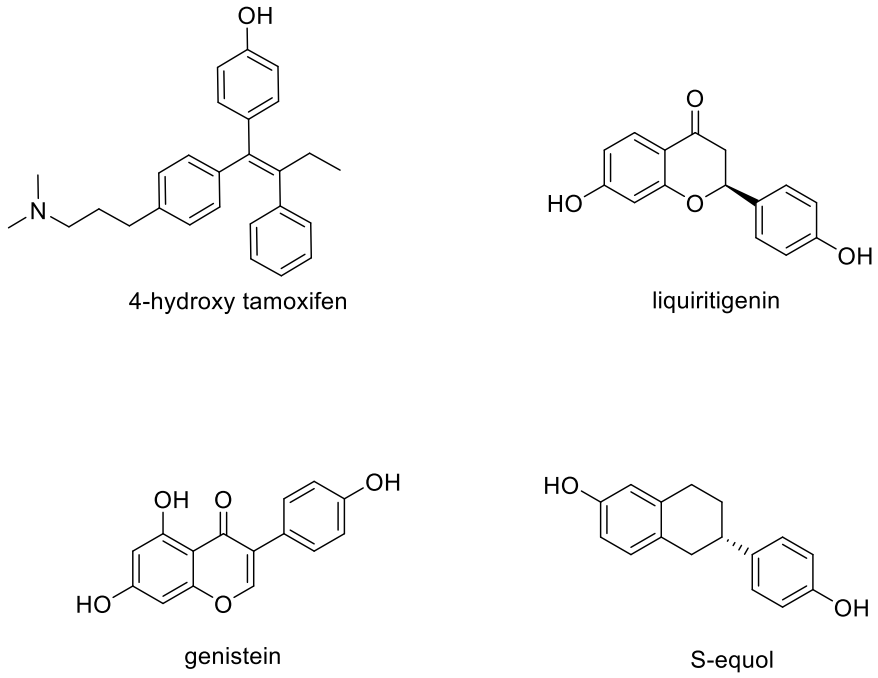


Figure 5.2 Known estrogen receptor β agonist. The above compounds are all known agonist for estrogen receptor β. While diverse structural commonalities can be identified through the use of phenol moieties and multiple aromatic cyclic systems.

5.1.2.1 Selective Ligands

Starting with Katzenebollgen's ER pharmacophore a more specific ERβ selective pharmacophore was generated [10]. The largest variance from the original

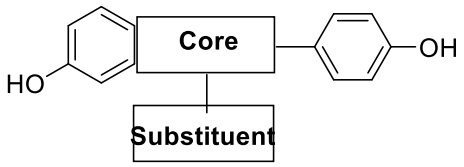


Figure 5.3 Estrogen Receptor β selective pharmacophore. The above pharmacophore was designed to generate isotype selective ligands.

pharmacophore is the fusion of the phenol ring with the central core. The core can serve as a focal point for selectivity due to the residue that sits above it. Of the two non-conserved residues Met-336 is located above where the C ring rest in ERβ binding pocket [11]. Though the conversion of Leu-384 to Met-336 is conservative in polarity, sterically the residues occupy different volumes [12]. The Met-336 residue allows for less steric hinderance than the Leu-384 (ERα) causing ERβ to forgo the linker space found in ERα pharmacophore. Using this model, a wide variety of agonist have been generated showing binding and selectivity with varying fused cores such as indazoles, benzothiophenes, benzofurans, indenes, phthalimides, indoles, imidazoles, naphthalene and quinoline [11]. While the selective pharmacophore embraced categories of compounds that were structurally different from our AMI class of ligands one certain group was of interest the arylindenones. Similar in structure to that of the AMI ligands, 2-Arylindene-1-ones overall showed lower ERβ selectivity however some showed good binding affinities for that of ERβ [13] The reported ERβ binding affinity for 2-arylindene-1-ones was attributed to the ability to maintain a hydrogen bonding network and the ligands size, both of which were criteria utilized in designing our ligands for ERα.

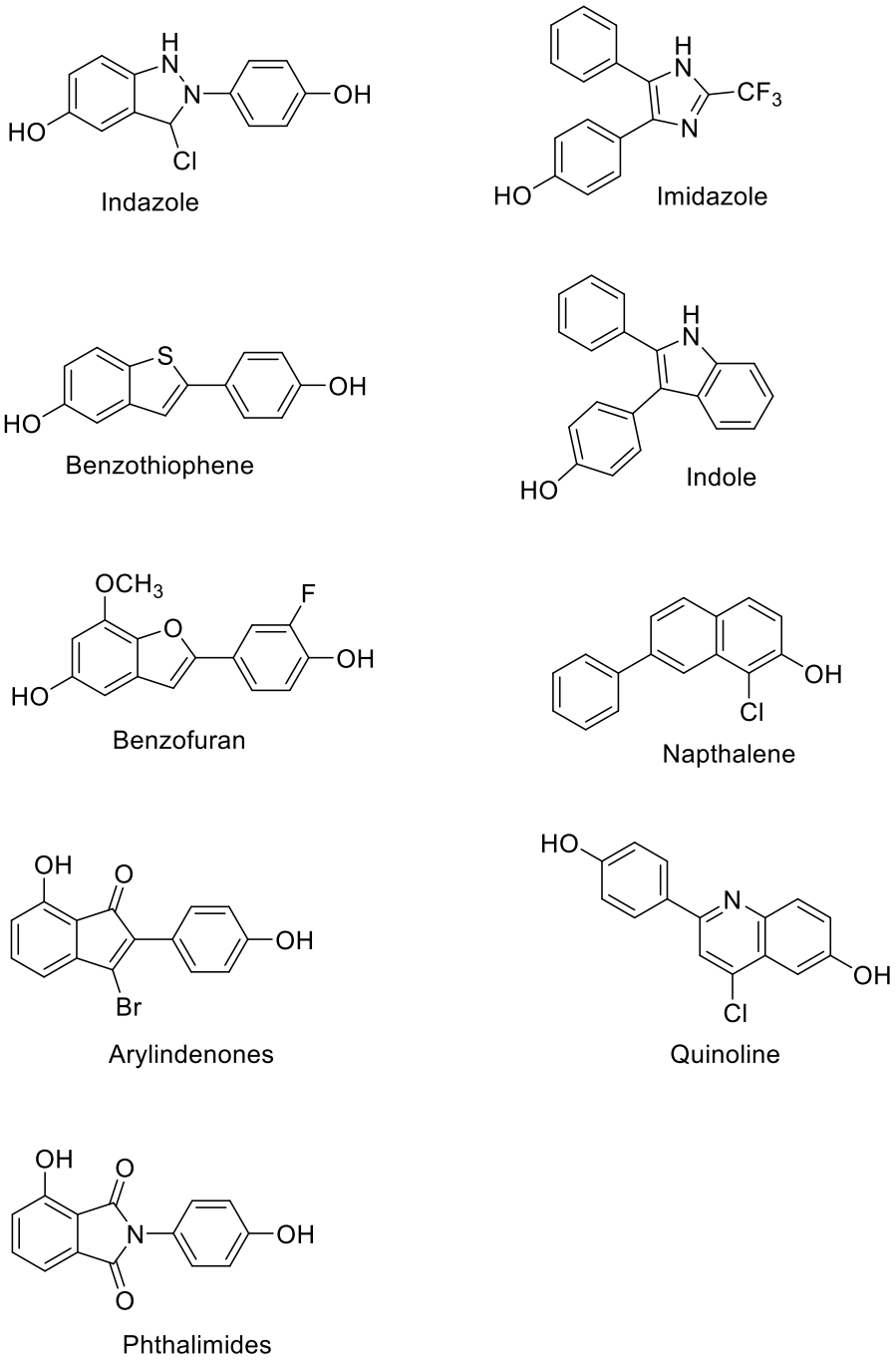


Figure 5.4 Core structures used in estrogen receptor β ligands. The above exemplifies the diversity in cores that can be used with generating ERβ ligands.

5.1.3 Physiological Role of ERβ

The physiological function of ERβ is much different from that of ERα with ERβ often directly counteracting the effects of ERα [14.]. Gene regulation of ERβ commonly leads to fluctuation of

cell cycle progression and apoptosis [15] which is opposite of its counterpart were activation typically promotes cell proliferation. Correlation between the loss of expression of ERβ has been well documented with breast, ovarian, and colon cancers where the progression of the cancer has been accompanied by suppression of ERβ [16] even completely absent in some cancers in more mature stages. Tumor suppression with higher levels of ERβ expression is suggestive that physiologically ERβ exerts a protective role against cancer development. The restraining ability of ERβ over its subtype counterpart has generated pharmaceutical interest for ligands that demonstrate specificity in activation. Exploitation of the physiological features of ERβ has been employed in exclusive targeting of the subtype in early stage of breast and colorectal cancers [17], however the complication resides in selective subtype agonist targeting and varying ratios of α/β expression. The presence of dual subtypes with opposing physiological response, varying distribution and expression levels underline both the need and complexity of ensuring that our designed ERα active ligands are target selective and function specific.

5.2 Design for Isotype Selectivity

Designing ligands with the capabilities to distinguish sites within the same class of nuclear receptor typically involves exploiting the differences between the target sites. While this approach can be used for designing ligands that are receptor family specific the strategy becomes more complex when selectively targeting subtypes such as ERβ vs ERα. In the case of these two subtypes the sole diverging features were the narrower pocket with smaller volume and the two differing residues in the pocket. Driven by our initial results of agonist binders we wanted to explore the influence of the ACD ring system and the pockets volume. Lastly, we wanted to examine what role if any chirality played in our AMI ligand design.

5.2.1 ACD

As outlined in chapter 4 our chosen core promoted an ACD ring system in addition to agreeing with the GFP chromophore structure the ring system agreed with the proposed nonselective estrogen receptor pharmacophore. Interestingly the proposed pharmacophore for subtype beta also encourages an ACD ring system.

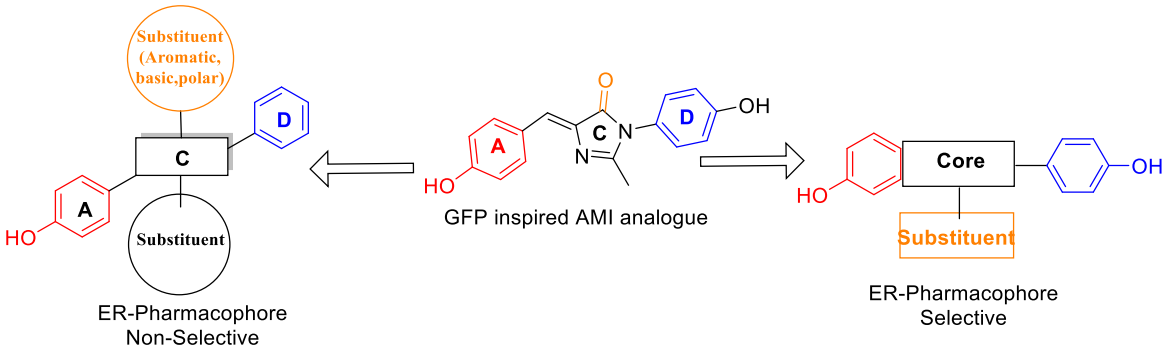


Figure 5.5 ACD ring system for both estrogen receptor selective and non-selective pharmacophore. Both pharmacophores utilize a ACD ring system like the GFP inspired ligand.

5.2.2 Chirality

Visual inspection of several estrogen receptor ligands identifies the presence of a stereocenter in a location where ligand-receptor residue interactions occur making it an optimal tool to guide selectivity.

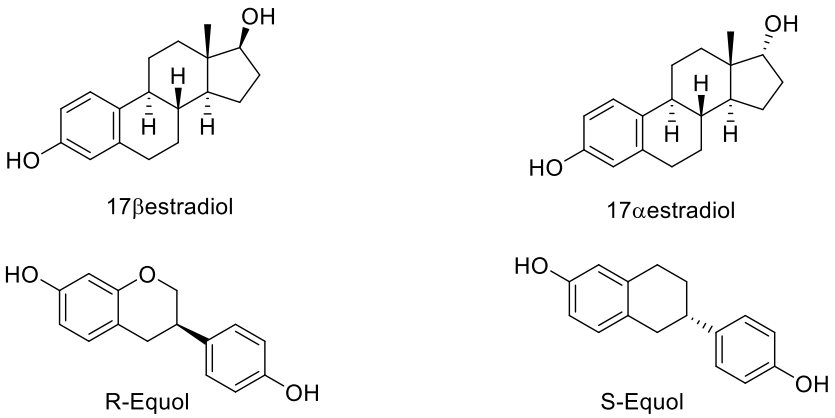


Figure 5.6 Estrogen receptor β ligands possessing chiral centers. The above ligands all bind ERβ and they all pose chiral centers in key locations that effect ligand-receptor interactions.

Natural ligands exist for ERβ that are selective or show preference such as S-equol [18] or17β-estradiol [19] and house a stereocenter that displays facial preference. The most commonly noted example of selectivity as a function of chirality is with the diarylpropionitrile or DPN enantiomers [20]. Of the two enantiomers the S-configuration possess the higher potency and selectivity for ERβ, the successful turn-on of transcription is also solely located in the S-configuration [21]. The variation in activity is attributed to the influence the stereocenter has on positioning the cyano group in contact with specific residues [22]. Exploration into the role that chirality could potentially play in our ligands promoted the synthesis of an additional set of AMI ligands. This 3rd generation of ligands were designed using A-rings with the hydroxyl in the para position and ortho position like the cores of CW34 and CW32 respectively as well as the 2,4-diol phenol seen in CW63. As illustrated in the figure below at the front of the ligand we incorporated a stereocenter bearing a hydroxyl group while maintaining a propyl chain. The importance of the hydroxyl group was to facilitate hydrogen bond acceptor interaction with the pocket but to see if a facial preference existed. The rationale for maintaining the propyl chain was to stay in compliance with ERα binding achieved. In the previous chapter we demonstrated that the presence of the small nonpolar substituent in front is beneficiary to ERα binding. Ideally incorporation of a stereo-center would promote selectivity by enhancing favorable residue contacts for ERα while minimizing access to key ERβ activating residues through establishing facial diversity.

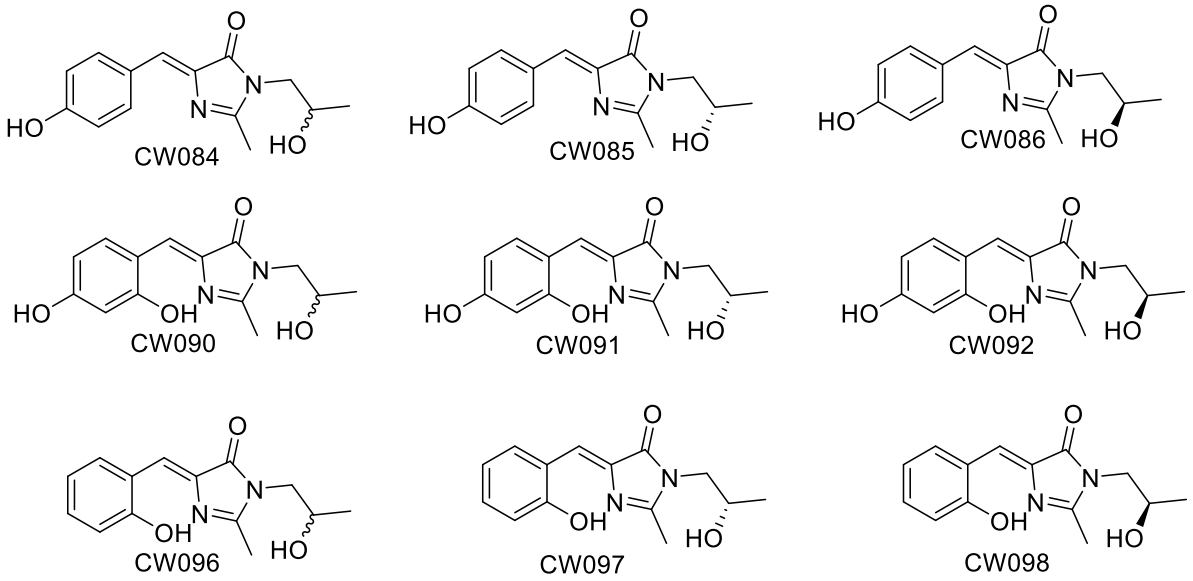


Figure 5.7 GFP inspired ligands generated to explore chirality. Ligands containing a chiral center were designed using template of the most active activator CW32.

5.3 Chemical Complementation

The same chemical complementation method was employed for the screening of ERβ with the difference being the use of ERβLBD:Gal4DBD fusion as opposed to ERα.

5.3.1 Estrogen Receptor α active ligands

The conservation of residues within the ligand binding domain of ERα and ERβ result in nondiscrimination of a vast number of ligands. In addition to ligands that produced a response, ligands of similar substitutions were also screened in hopes of possibly discovering a lead for the same class of ligands that consequently only activate ERβ. Although ultimately all compounds were screened against ERβ of primary interest was molecules that showed activity in ERα assays. Once again growth was observed for CW32 only. Optical density levels for CW32 were observed at 0.60 ± 0.05. Identical controls were employed with NL demonstrating an OD of 0.04 ± 0.02 and estradiol an OD of 0.48 ± 0.01, in addition the same positive control (Gal 4) was used

expressing OD values around 0.61 ± 0.1. The reduced optical density value of CW32 in ERβ may imply that a slight preference for ERα is possible like the profile of estradiol with ERα and ERβ. Overall, activation as observed via chemical complementation for selective estrogen receptor activation indicates that CW32 activates both ERα and ERβ subtypes however possibly with slight partiality.

Table 5.1 Estrogen receptor β chemical complementation.

Ligands	Optical Density
CW32	0.60 ± 0.1
Ethynyl Estradiol (EE2)	0.50 ± 0.3
Estradiol (E2)	0.48 ± 0.01
Gal 4 (positive control)	0.63 ± 0.1
No Ligand (NL)	0.04 ± 0.02

5.3.2 AMI Library Screen

In addition to the ERα active ligands the remainder of the compounds generated through out 1st and 2nd generation synthesis was screened. No additional ligands from these two generations produced results that would be considered significant. Chemical complementation screening of AMI library resulted in no observation of activation for ERβ.

5.3.3 Generation III Results (Chiral AMI Ligands)

Arylmethylene imidazolone ligands containing a chiral center were screened as third generation ligands exploring chirality. Since the incorporation of a chiral center did not occur until

investigation of receptor subtype selectivity third generation ligands were screened for activity in ERα as well as ERβ. Like screening of previous generation controls included the no ligand (NL) vehicle, Gal 4 and estradiol. No growth was observed in the NL yielding an OD of 0.04 ± 0.01, positive controls Gal4 and estradiol gave an OD's of 0.665 ± 0.2 and 0.454 ± 0.1 respectively. Chemical complementation screening of 3rd generation ligands resulted in no observation of activation.

5.4 Luciferin Reporter Assay Analysis for Isotype Specificity

Using the same methodology employed for ERα screening, the LBD of ERβ was fused to Gal4 DBD and transformed into pCMX mammalian vector. Identical to ERα analysis compounds were screened in HEK293T cells and estradiol was employed as a positive control. It is important to note here that estradiol exhibits a preference for ERα over that of ERβ here go it is expected that observed control values for ERβ will be lower. Estradiol's observable accepted value for our ERβ analysis is 14 ± 0.2 as opposed to 26 ± 0.7 in ERα.

5.4.1 Estrogen Receptor α Active Ligands

Initial screening was conducted to assess compounds that have shown activation in ERα. Of the set of activators CW32 displayed considerable activation above baseline 6.4 ± 0.3 almost half that of estradiol. Results indicate that ERβ activation for CW32 is far less sensitive to small structural modifications than that of its counterpart. Rotation of the phenol from the ortho position to the para position CW28 results in loss of activity with a reading just above baseline at 2.5 ±0.04. Likewise, alteration of the amide substituent from isopropyl to propyl CW32 and CW30 also resulted in loss of all activity 1.6 ± 0.1. Abolishment of activity was also observed with CW35 and CW38 wherein the isopropyl group was converted to cyclopentyl and

cyclohexyl rings yielding results of 2.0 ± 0.2 and 2.0 ± 0.24 respectively. Of the compounds initially designed to assess the role of hydrogen bonding in the front of the pocket only CW34 showed activity above the baseline at 2.8 ± 0.02 roughly a third that of CW32. The results from the remainder of the ligands can be viewed in table 5.8

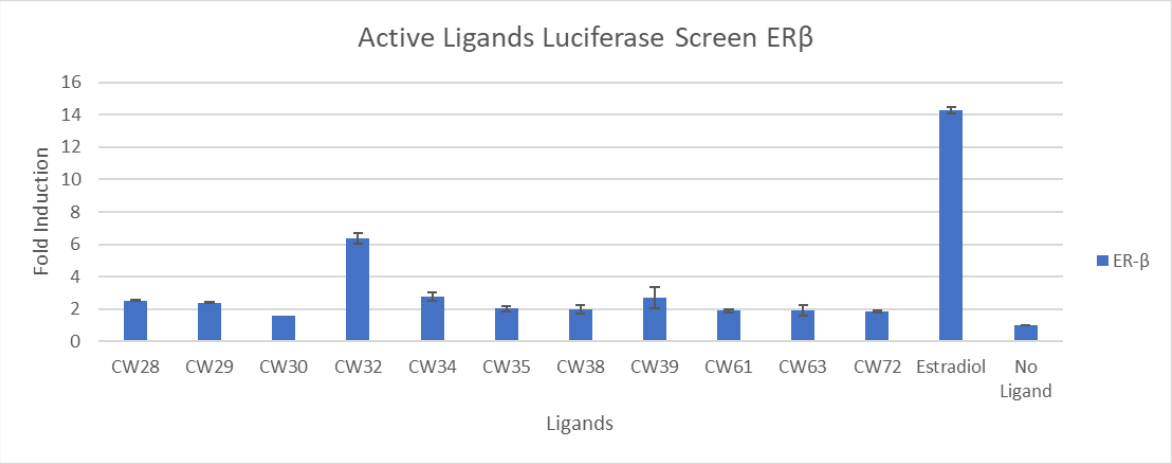


Figure 5.8 Screening of ERα active ligands against ERβ. Ligands shown to be active in ERα were tested against ERβ for selectivity.

Table 5.2 Results of ERα active ligands screened against ERβ.

Ligands	Fold Induction
CW28	2.5 ± 0.04
CW29	2.4 ± 0.1
CW30	1.6 ± 0.04
CW32	6.4 ± 0.3
CW34	2.8 ± 0.3
CW35	2.0 ± 0.2
CW38	2.0 ± 0.2
CW39	2.7 ± 0.6

CW61	2.0 ± 0.9
CW63	1.9 ± 0.3
CW72	1.8 ± 0.1
Estradiol (E2)	14 ± 0.2
No Ligand (NL)	1

5.4.2 AMI Library Screen

In addition to the ERα active ligands the remainder of the compounds generated through out 1st and 2nd generation synthesis was screened. No additional ligands from these two generations produced results that would be considered significant.

5.4.3 Generation III Results

Chiral ligands were screened including the racemate and the separate enatiomers, as stated previously these ligands were not generated during the initial ERα screen therefore in addition to ERβ, ligands were screened against ERα as well. Surprisingly no 3rd generation ligands produced activation higher than that of baseline for ERα or ERβ. Worth considerable mention are the results of CW86 in ERβ in comparison with the same ligand in ERα. While CW86 produced normal baseline activity in ERα (1.5 ± 0.3) with ERβ baseline activity suppressed with a reading of 0.87 ± 0.03. While exploration into estrogen receptor subtype selective antagonist is beyond the motivation of this dissertation this compound may prove as a starting point for future work. Also, worth noting is ligand CW98 that showed response slightly above that of baseline for ERα (2.7 ± 0.1) and a normal baseline for ERβ (1.5 ± 0.2).

5.5 Time Resolved Fluorescence Resonance Energy Transfer (TR-FRET)

This assay employs estrogen receptor α and β labeled via a GST tag bound to its anti-GST antibody that is terbium labeled. Competitive binding to estrogen receptor occurs via our AMI agonist ability to displace a fluorescent ligand in our case fluorescently labeled estradiol (ES2 Green tracer). The proximity of labeled ligand and the tagged antibody directly corresponds to the strength of the FRET signal [23]. Here successful displacement of the tracer by our AMI agonist will result in a decrease in FRET signal. This experiment was conducted to strengthen our hypothesis that all binding observed by our agonist occurred in the same pocket as that of estradiol and not via any alternative binding site. This experiment was also carried out to support the result of ERβ screening which demonstrated preference for subtype with select ERα active agonist. A subset of four compounds were chosen to be tested in ERα and ERβ. The compounds selected included the highest two activators CW32 and CW34 as well as the lowest activator CW72, lastly as a negative control CW33 was used which had proven to be a non-binder in all assays. As shown in the charts below our highest two activating compounds displace estradiol in both ERα and ERβ, however both with preference to ERα. While the lowest activator in our assay agonist CW72 exclusively displaced estradiol in ERα and displayed no displacement in ERβ. Table 5.3 list displacement percentages of tested compounds for both estrogen receptor subtypes.

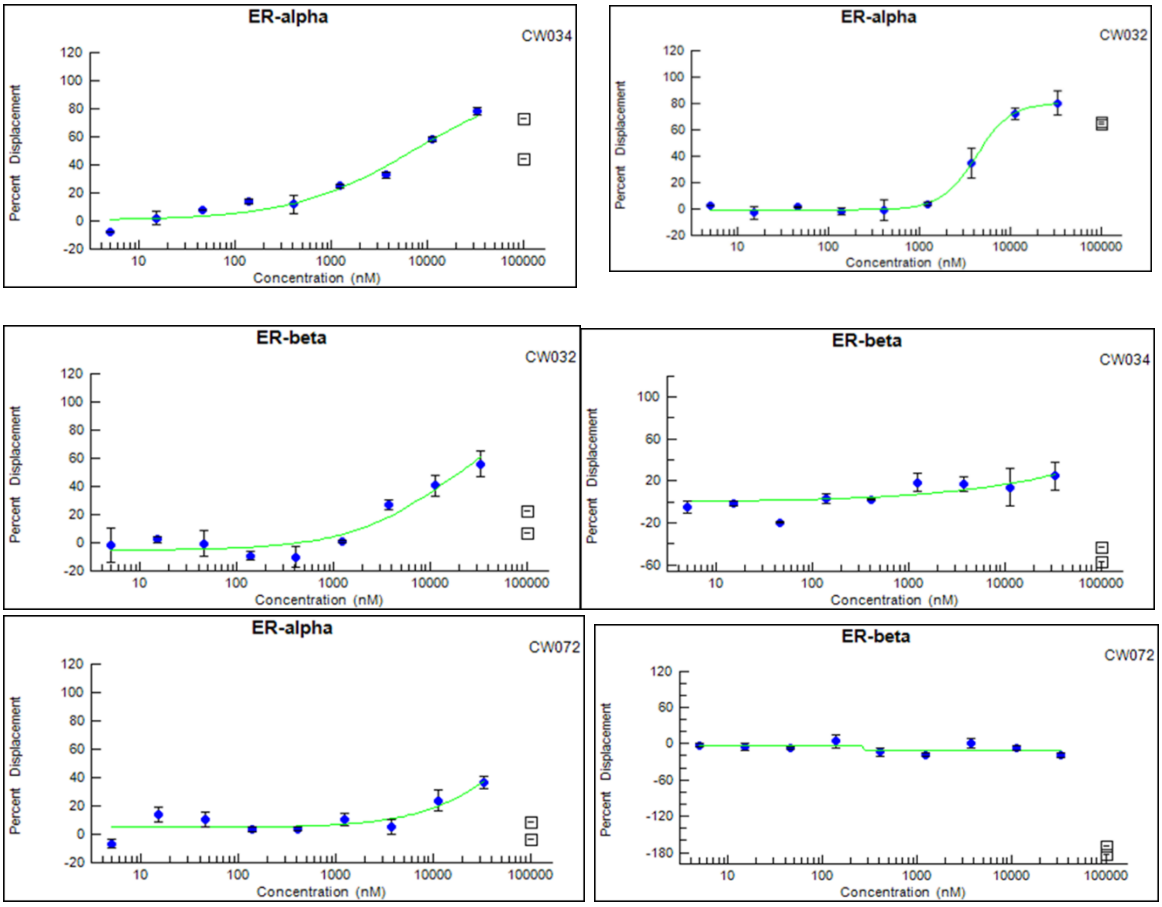


Figure 5.9 TFRET dose response curve for select GFP-inspired ligands. Select ligands were chosen (top 2 activators, lowest activator and a non-activator) and assayed via TFERT assay the 10pt titration curve for both isotypes are illustrated.

Table 5.3 Percentage of estradiol displaced by select GFP-inspired estrogen receptor agonist. Select ligands were chosen (top 2 activators, lowest activator and a non-activator) and assayed via TRFRET assay the percentages of the displaced tagged indigenous ligand are listed for both isotypes.

Ligands	ERα	ERβ
CW32	89%	55%
CW34	75%	24%
CW72	37%	0%
CW33	0%(non-activator)	0% (non-activator)

5.6 Selective Ligand Activation Discussion Rationale ERβ

As stated previously differential targeting in the ligand binding pocket between ERα and ERβ can be complex due to only two substitutions in amino acid residues within the pocket. While the variation in pocket size contributes to ligand selectivity certain factors are still necessary for successful binding within the pocket. These criteria are the same as what we established in the designing of our ligands recall attention was given to: 1) preservation of hydrogen bonding network via key amino acid residues within the ligand binding pocket 2) size and volume occupied by the ligand and lastly 3) molecular landscape of the ligand with respect to polarity therefore, analysis of results from ERβ was developed from the above criteria and explained through those guides.

5.6.1 Hydrogen Bonding Network

With key residues such as arginine, glutamic acid, and histidine unchanged within the ligand binding pocket the opportunity for ligands to participate in a network of hydrogen bonding with the receptor is parallel to that of its counterpart.

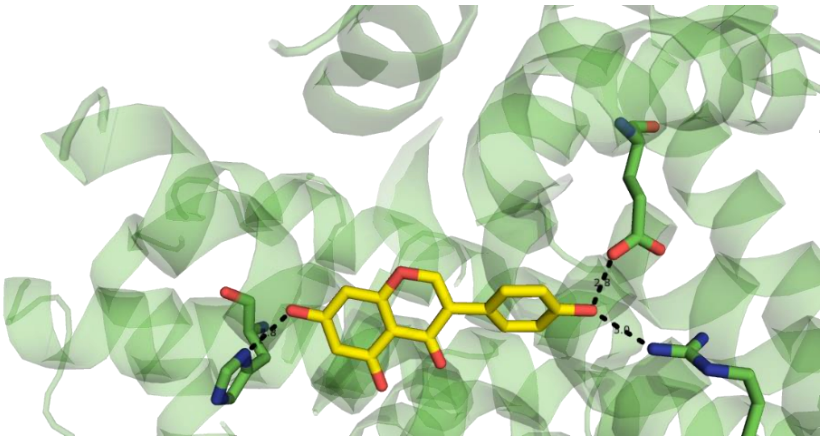


Figure 5.10 Hydrogen bonding network within estrogen receptor β ligand binding pocket. Key residues shown as sticks involved in hydrogen bonding network with selective agonist genistein shown as red (oxygen) and yellow (carbon) sticks.

The preservation of the hydrogen bonding network is exemplified through Richardson etal. studies on benzopyrans as selective estrogen receptor β agonist or SERBA [24-27]. In the case of the benzopyrans the incorporation of two phenols located at the ends of the ligand analogous to A and D rings in estradiol resulted in water assisted hydrogen bonding with Gln-305, Arg-346 and direct hydrogen bonding with His-475. Ligand modeling in chapter 4 illustrated that our ERα active ligands participate in hydrogen bonding either with the 3 key residues or with the carbonyl backbone in amino acid residues. As illustrated in figure 5.10 the same opportunities exist with ERβ. Figure 5.11 shows a significant increase in the distance of ligand and amino acid residues involved in hydrogen bonding.

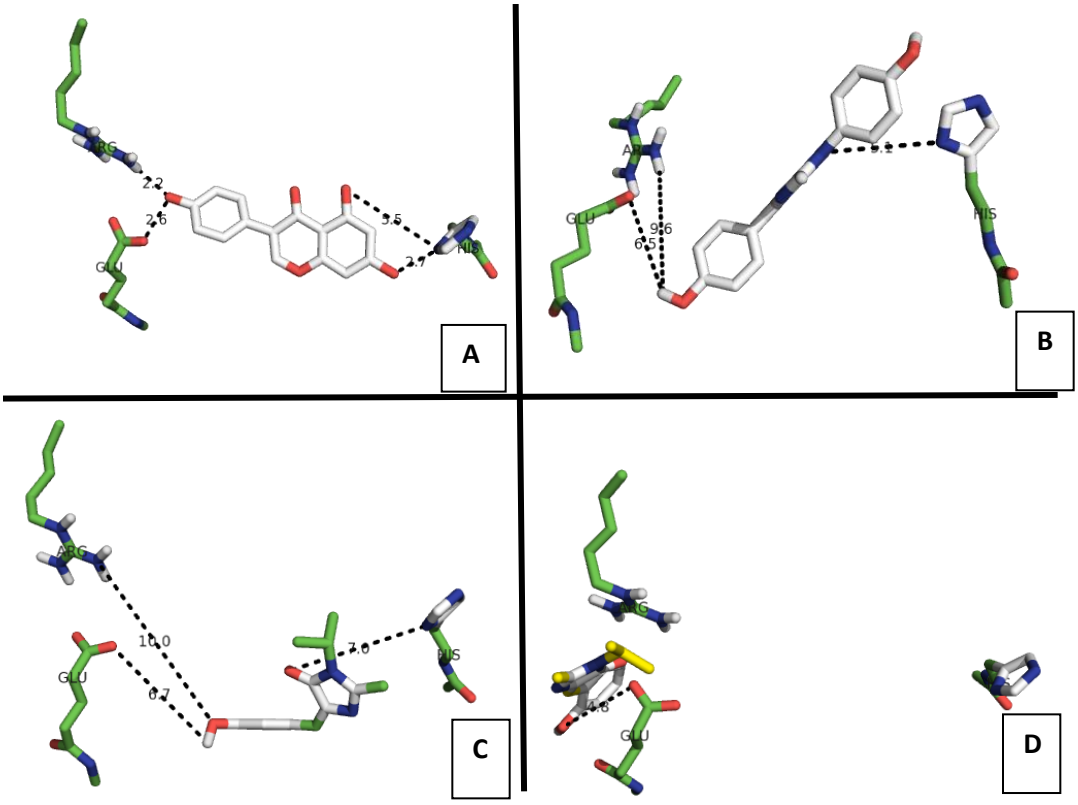


Figure 5.11 Hydrogen bonding network within ERβ. **A**) Genistein (shown in middle) participates in hydrogen bonding with all 3 key residues as illustrated all with a distance under 4A (shown as dashed lines in black). **B**) CW34 can hydrogen bond with all 3 key residues however distances are well over 4A. **C**) CW28 can participate in hydrogen bonding with 3 key residues however

distances as shown with black dashes exceed that of 4A. **D**) CW32 only interacts with Gln-305 however does not interact with the other 2 key residues.

The large distance suggest that hydrogen bonding is not likely to occur even if it were water mediated. CW32 while it does not interact with all 3 key residues in the ligand binding pocket it does participate in hydrogen bonding relationship with Gln-305 and is under 4 A distance. It appears the main reason successful hydrogen bonding is not achieved is due to the placement of the ligand in relation to the LBP of ERβ. In comparison to its predecessor where the priority was centered around the ability to participate in a hydrogen bonding network it appears here priority must be given to ligand volume to ensure additional criteria can be supported.

5.6.2 Ligand Volume-Size Shape

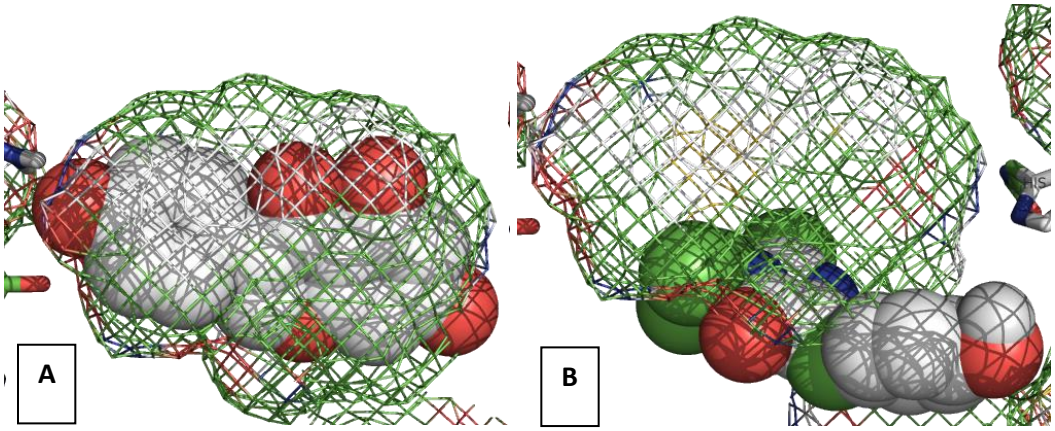
The smaller size of ERβ ligand binding pocket is the most distinguishing difference of the two binding pockets and pose possibly the greatest restrictor on selectivity and activation. Activators of ERα that are considered rather large such as the tetra substituted propylpyrazole (PPT) or the triarylamide's are excellent agonist for ERα however display minimal affinity for ERβ. Comparisons of molecular volumes of several known ERβ agonist with select AMI ligands are illustrated in Table 5.4 indicate relatively comparable size.

Table 5.4 Molecular volume and selectivity for known estrogen receptor β ligands and GFP inspired AMI ligands.

Ligand	Volumes (Angstroms)	Isotype Preference
DPN	222	ERβ
Genistein	272	ERβ

Benzoxazole	226	ERβ
Isocoumarin	233	ERβ
Isobutestrol	222	ERβ
CW32	230	ERα
CW33	222	Neither
CW34	259	ERα

While the volume of both CW32 and CW34 are well within the limits of acceptable agonist such as genistein and highly selective DPN; CW34 and CW32 orientation within the ligand bonding pocket of ERβ presumably contributes to its preference for ERα. As molded in figure 5.12 genistein is oriented within the pocket however the majority of the GFP inspired ligands do not align with in the pocket, CW34 even appears to be larger than the binding pocket.



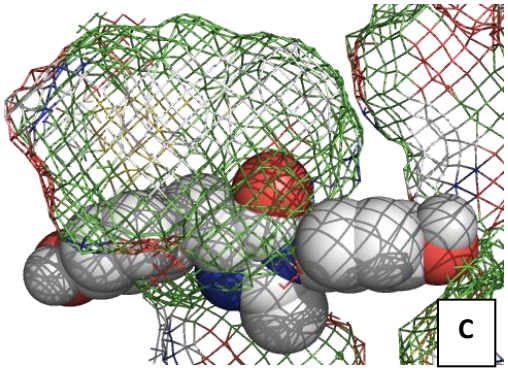


Figure 5.12 Modeling of volumetric space of Estrogen receptor β ligand binding pocket and select ligands. **A**) Genistein shown as a sphere inside of the estrogen receptor ligand binding pocket shown in mesh. **B**) CW28 shown as spheres sits slightly under the ligand binding pocket with portion of the ligand in the pocket (mesh). **C**) CW34 shown as spheres indicates that only the middle of the ligand sits inside of the ligand binding pocket (mesh) while the remainder of the ligand rest outside the pocket.

Activators of ERβ tend to be slenderer to accommodate the smaller sized pocket, ligands such as diarylpropionitrile (DPN) and isobutestrol exemplify this. Studies show that as size of the ligand increases activation may increase however selectivity diminishes [28]. In the case of ligand CW32 modeling indicates that the ligand does not exist within the ligand binding pocket however resides just to the left in what appears to be a secondary pocket as shown in figure 5.13 below.

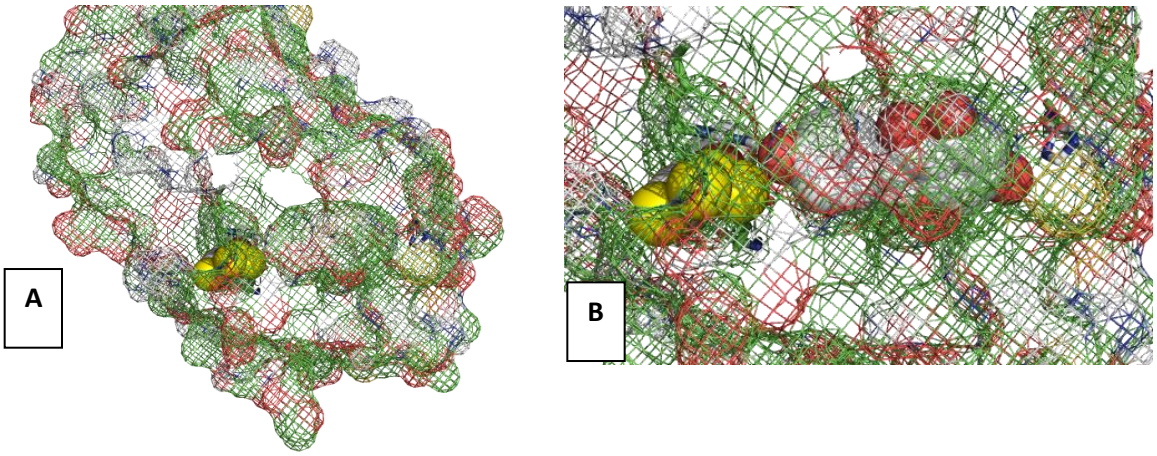


Figure 5.13 Placement of CW32 in secondary pocket within ERβ. **A**) CW32 shown as yellow spheres sitting in pocket adjacent to the primary ligand binding pocket. **B**) CW32 shown as

yellow spheres resting in an additional binding pocket while genistein shown as spheres resides in the ligand binding pocket.

5.6.3 Molecular Topology

Though drastically similar in amino acid residue the molecular surface area of the ligand can have a large impact on binding and activation. Ligands that are known to be selective or potent for estrogen receptor β were computationally calculated to have varying polar surface areas suggesting that ERβ ligand binding pocket is tolerant of diverse polarity within ligands. Table 5.5 below list various AMI ligands with varying cores in comparison to known binders. In addition to the polar surface area of the ligand the landscape of the ligand binding domain is important as well. Analysis of table 5.5 indicates that while polar surface area may be a criterion adhered to for overall estrogen receptor binding and activation it does not appear to have bearing on selectivity.

Table 5.5 Polar surface area of select AMI ligands and ERβ selective ligands.

Ligand	Polar Surface Area
DPN	64
CW32	55
CW34	75
Hexestrol	40
Genistein	91
Benzothiophene	40
Indazole	56
Estradiol	40

5.6.4 Chirality

The incorporation of a chiral center to enhance selectivity for estrogen receptor α did not prove to be useful as the ligands possessing chirality were not active.

5.7 Identification of Active Leads for Estrogen Receptor

Summation of chemical complementation and luciferin reporter gene assay results revealed a series of substituted arylmethyleneimidazolones that are viable as nuclear receptor binders. Further exploration demonstrated that several of these analogues serve as estrogen receptor agonist bearing specificity or strong preference for subtype α as opposed to β. While analogues CW72 was completely exclusive for estrogen receptor α others such as CW32 exhibited activation in both isotypes α and β however displayed a preference for ERα. Activation intensity varied from levels well below estradiol, to comparable and even exceeding estradiol. The robust number of compounds synthesized and evaluated allowed for creation of a suitable structure activity relationship analysis which can be fitted to known activation rationale parameters.

The generation of several estrogen receptor agonist along with identification of subtype preference prompted the next aim, investigating selectivity on a boarder scale to ensure that activating compounds were nuclear receptor selective otherwise known as SNuRM. Progressing forward only compounds listed as estrogen receptor α binders (Chapter 4) resulting in agonist activity will be the basis of exploration providing further depth to selectivity of binding compounds.

5.8 Nuclear Receptor Selectivity

The presence of additional families of nuclear receptors in the body creates alternative opportunities for binding of ligands. Nuclear receptors domains are held constant however diversity is demonstrated largely across families especially within the ligand binding domain and even more in the ligand binding pocket however it is possible for ligands that activate one family to possess favorable properties for another. While this may seem like a positive side effect specifically for receptors that work in concert where activation or repression of both yield an enhanced desired outcome it can be disastrous for receptors that turn on unwanted physiological responses linked to diseases. This rationale prompted the screening of only active ligands to ensure that selectivity can be observed across nuclear receptors with the implication more selective ligands should produce minimal undesired side effects. Chemical complementation screening as outlined in chapter 4 was employed in assessment of compounds identified as agonist for estrogen receptor α. Both retinoic acid receptor γ (RARγ) and rexinoid x receptor α (RXRα) ligand binding domain where used in place of estrogen receptors ligand binding domain.

5.8.1 Rexinoid X Receptor α

Rexinoid X Receptor (RXR) is an intriguing player within the nuclear receptor family while debate exist as to the physiological role of the receptor, RXR is highly coveted for its ability to form heterodimers with other nuclear receptors. Partners for the RXR heterodimerization include LXRα, PPARα, CARβ, PXR and FXR all of which have been reported that in the absence of the ability to heterodimerize with RXR the physiological pathways controlled by those receptors are compromised [29]. Heterodimers generated by RXR exist as one of two categories either

permissive such as PPAR, LXR, PXR or non-permissive such as RAR, VDR, TR [30, 31]. The distinction between the two categories are that for permissive partners binding of an RXR ligand co-promotes ligand binding of the partner receptor while in non-permissive heterodimer pairs RXR serves as a silent partner primarily functioning to assist in binding of the hormone response element. With the ability to form combinatorial partners with one third of the 48 human nuclear receptors RXR represents an alternative to activating multiple signaling pathways and an important receptor to ensure that non-targeted RXR ligands do not activate.

Rexinoid x receptor exist in three isotypes α, β, γ attention in literature has been given to RXRα [32] therefore focus for selectivity will center on RXRα. Isotype RXRα is abundantly expressed in the liver, kidney, spleen and placenta epidermis as well as a variety of visceral tissues [33,34]. Reminiscent of the role over expression of ERα plays in breast cancer, over expression of RXRα has been reported in 66% of breast ductal carcinomas [35]. Structurally composed of 462 amino acid residues it is smaller than that of ERα [34]. The LBD is composed of the common three-layered alpha-helical sandwich and at 238 amino acid residues it is less than 20 amino acid residues smaller than ERα LBD (253).

5.8.1.1 Ligand Binding Pocket Comparison

Through a comparative analysis of the ligand binding pocket of RXRα with that of ERα a clear understanding of how to promote selectivity in binding can be obtained. Architecturally the ligand binding pocket of RXRα is 400-500A [34] comparable to the volume of ERα (450A). With such comparable volumes spatially, ligands that can accommodate ERα ideally can fit into RXRα LBP as well. Comparing the type of amino acid residues that reside within both pocket as shown in figure 5.14 RXR ligand binding pocket does poses amino acid residues arginine, glutamine and histidine in an orientation reminiscent to that of ERα.

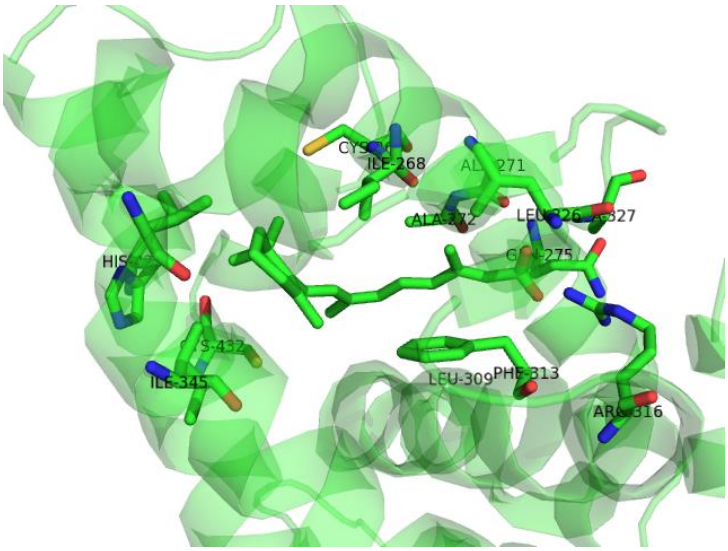


Figure 5.14 RXRα ligand binding pocket complexed with 9-cis retinoic acid. The LBP of RXRα different from ERα however contains key residues Histidine, Arginine and Glutamine that can establish a hydrogen bonding network.

While the ligand binding pocket of ERα and RXRα host diverse ligands in structure lithocholic acid an agonist for RXRα contains the same ABCD skeletal ring system as that of estradiol the LBP can also accommodate ABD ring systems figure 5.15. Of even more significant commonality are the use of the hydroxyl moiety on the A-ring and the hydrogen bond donating carboxylic acid on the D-ring recall both of which advocate binding in ERα. Given these overlaps consideration must be given to the possibility that estrogen receptors pharmacophore may produce suitable ligands for RXR.

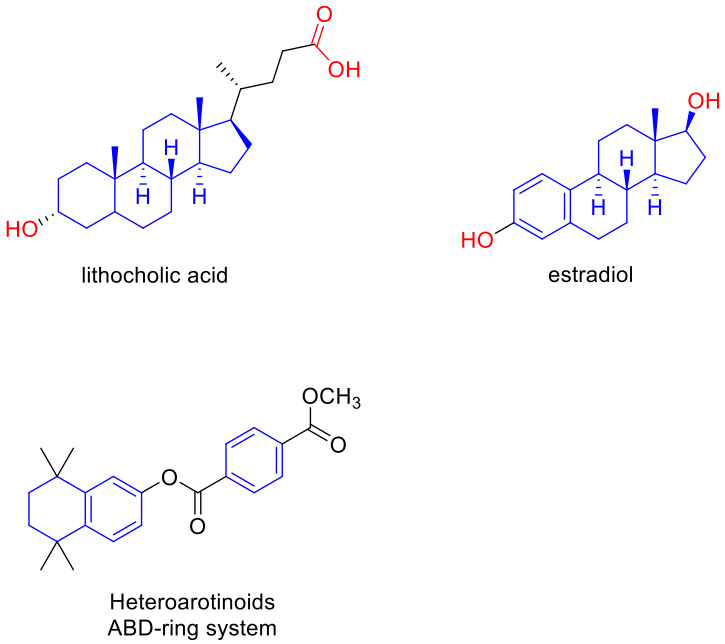


Figure 5.15 RXRα binders with similar ring systems and functionalities as ERα agonist. Ligands capable of binding RXR are illustrated with ring system (blue) favorable for ERα. In the case of lithocholic acid favorable functionality (red) is present as well. The structural of estradiol ring system (blue) and functionality (red) is presented for comparison.

5.8.1.2 Rexinoid X Receptor Chemical Complementation Results.

Active estrogen receptor AMI ligands based on the summation of aim 1 were evaluated for activation of rexinoid x receptor using chemical complementation. The experimentation used 9-cis retinoic acid as the positive control ligand and exhibited an optical density (OD) of 0.3 ± 0.04. Controls in the form of no ligand (NL), and Gal 4 identical to previous chemical complementation screening were employed. No growth was observed in the NL yielding an OD of 0.04 ± 0.01 and the positive control Gal4 gave an OD of 0.5 ± 0.1. Chemical complementation screening of active ligands resulted in no observation of activation from any test ligands.

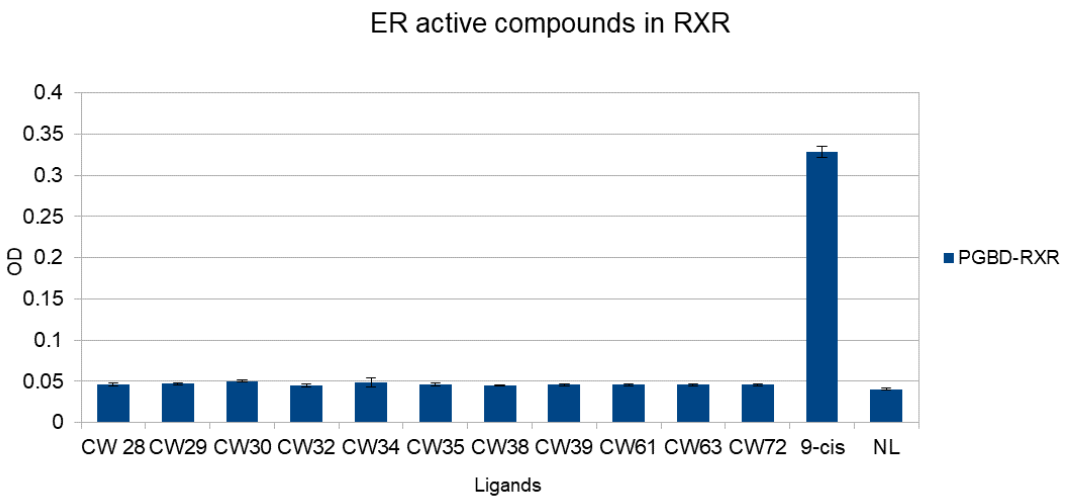


Figure 5.16 Chemical complementation of ER active ligands against RXRα. Chart above indicates activation in only 9-cis retinoic acid (control).

5.8.2 Retinoic Acid Receptor γ

Retinoic Acid Receptors or RAR chiefly forms heterodimers with RXR as the partner however ligand mediated activation is achieved through the binding of RAR [30]. Focus on retinoid ligand development has led to several approved drugs for this family of receptors coming to market. Occurring in three isotypes α, β, γ priority was assigned to RARγ as it is the most divergent of the isotypes [36]. Composed of 454 amino acid residues RARγ is smaller in size than ERα however the implications of the RAR family in cellular differentiation, proliferation and apoptosis outlines a role in cancer like that of estrogen receptors [37]. Specifically, in skin cancer RARγ is a predominate player and has 90% expression in epidermis studies have indicated that decrease in the levels of expression coincides with tumor progression [38].

5.8.2.1 Ligand Binding Pocket Comparison

Comparison of LBP for RARγ and ERα display distinct difference. The largest contrast being the presence of residues necessary to establish a hydrogen bonding network akin to ERα. As

illustrated in figure 5.17 of RARγ LBP coupled with 9-cisretionic acid, an arginine residue is present however histidine and glutamine are not. It is worth noting that a serine residue is present and capable of hydrogen bonding. The LBP of RARγ has been described as a I-shaped pocket with a bulge accommodating ligands with larger steric bulk [30] this contrast in pocket spacing could aide in selective ligand recognition.

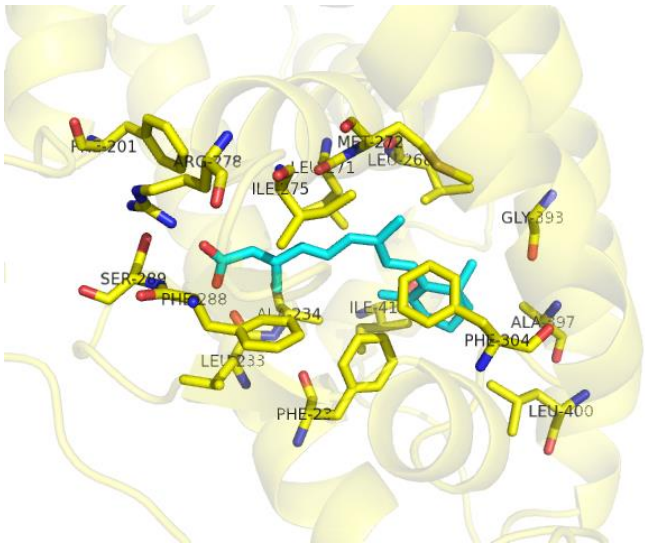


Figure 5.17 Ligand binding pocket of RARγ. Key residues of RARγ ligand binding pocket (sticks) are illustrated with its interactions to 9-cis retinoic acid (green sticks).

5.8.2.2 Chemical Complementation Results

Active estrogen receptor AMI ligands based on the summation of aim 1 were evaluated for activation of retinoic acid receptor γ using chemical complementation. The experimentation used all-trans retinoic acid as the positive control ligand and exhibited an optical density (OD) of 0.36 ± 0.1. Controls in the form of vehicle (ethanol:dmso), no ligand (NL), and Gal 4 identical to previous chemical complementation screening were employed. No growth was observed in the NL yielding an OD of 0.04 ± 0.02 and the positive control Gal4 gave an OD of 0.6 ± 0.2

Chemical complementation screening of active ligands resulted in no observation of activation from any test ligands.

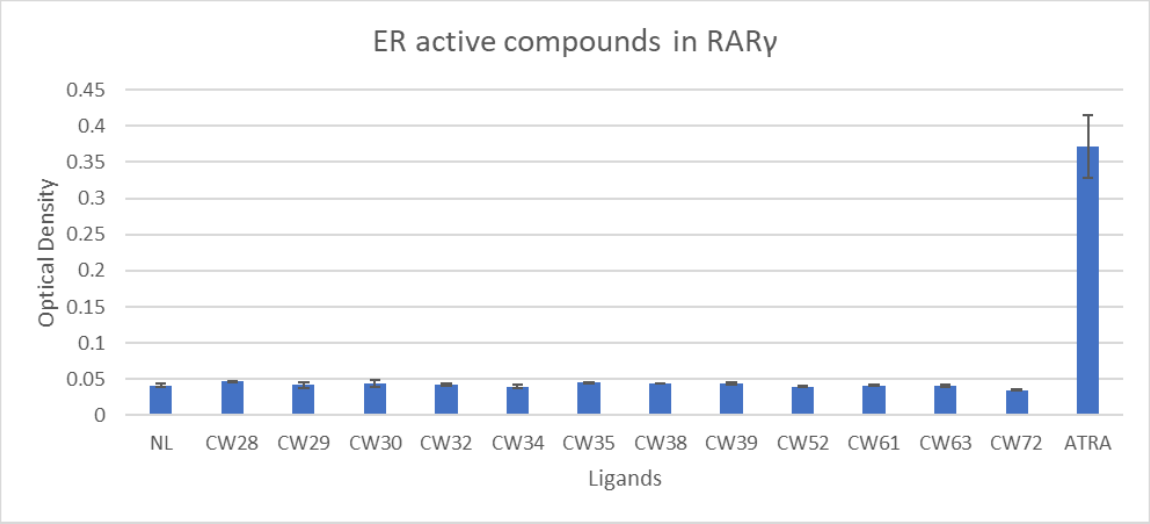


Figure 5.18 Chemical complementation for estrogen receptor active compounds against RARγ. Control all trans retinoic acid was the only compound that displayed activation.

5.9 SUMMARY

The combination of nuclear receptors ability to exist as isotypes and bind similar classes of ligands emphasis the importance of developing ligands that are both isotype and nuclear receptor selective. Here in we have demonstrated that a select set of ligands shown to be active in ERα can be selective for its isotype ERβ and in the case of CW72 can be exclusive. The highest two activating ligands CW32 and CW34 demonstrate good disparity in preference for ERα over that of ERβ. Through modeling the rationale for selectivity can be rooted in the absence of hydrogen bonding and improper orientation in the ligand binding pocket not satisfying two of the three criteria by which the ligands designed centered around. The introduction of a chiral center to influence molecular topology proved to be unsuccessful. While ERα ligands exhibited isotype

preference globally the ligands displayed nuclear receptor selectivity demonstrating no activation in RXRα or RARγ.

5.10 Materials and Methods

Chemical Complementation in Yeast

Conventionally chemical complementation is genetic selection highlighted by the use of a small molecule. Herein we use a three-component system linking the survival of yeast to the ability of a small molecule modulator to activate nuclear receptor transcription. The yeast strain used (Saccharomyces cerevisiae – PJ69-4A) contains Gal 4 response elements (Gal4RE) used to control the expression of genetic selection genes involved in the histidine (HIS3) and adenine (ADE2) biosynthetic pathway. Transformed in the yeast are two fusion proteins, a Gal 4 DNA binding domain (Gal4DBD) fused to a nuclear receptor ligand binding domain (LBD) and a Gal 4 activation domain (GAD) fused to a nuclear receptor co-activator (SRC-1). At the core of chemical complementation binding of a ligand to the nuclear receptor LBD promotes binding of the LBD: Gal4DBD fusion with the Gal4RE thus prompting a conformation change of the receptor allowing for recruitment of the GAD:SRC-1 fusion. Completion of this cascade actuates transcription of selective genes (HIS3 or ADE2) thereby allowing yeast to survive in media lacking the presence of histidine or adenine. (Fig.3). Yeast trans formant (pGBDhERαLBD) was grown overnight in 2mL of media lacking leucine and tryptophan (SC-LW) at 30°C with shaking at 300rpm. After growth, yeast was pelleted (4°C, 7min, 3700rpm,) media discarded and re-suspended in 1mL of water. A second pellet was formed and re-suspend in 1ml of water before a 45µL aliquot was added to 5mL of water. Media lacking histidine, leucine and tryptophan (SC-

HLW with 0.1mM 3AT) was prepared and ligand-media mixes were made via dilutions from pre-made ligand stock solutions ($10^{\wedge -2}$ M) with a final concentration of 10µM. The incorporation of 3AT (0.1mM) was necessary to suppress background arising from imidazole glycerol-phosphate dehydratase a key enzyme in the histidine biosynthetic pathway. In the case of dose responses serial dilutions were performed starting with 10 µM, solutions ranged from (10^{-5}M-10^{-12}M). Using a multi-channel pipetor 20µL of cells were added to 80µL of ligand in a 96 well plate. To wells not containing cell: ligand mix, 80µL of water was added. Plates were incubated 48hrs (30°C and 300rpm) with optical density (OD) measurements recorded at 0, 24, 48 hour intervals. Optical density readings were taken at a wavelength of 630 nm.

Mammalian Cell Culture Assay

Human embryonic kidney 293T (HEK293T) cells (ATCC, USA) were transfected with the plasmid pCMXwithERαLBD. This plasmid contains the Gal4DBD (GBD) fused to the wild-type ERα ligand binding domain LBD (GBD:LBD fusion under the control of a cytomegalovirus (CMV) promoter). The reporter plasmid, p17*4TATAluc, contained the Renilla luciferase gene under the control of four Gal4 response elements located upstream from a minimal thymidine kinase promoter. The pCMXβgal, a plasmid containing the β-galactosidase gene under the control of the mammalian CMV promoter, was also used as an internal standard. Lipofectamine 2000 (Invitrogen, USA) was used as the cationic lipid and transfection experimental details are described in Taylor et al. Ligands were added to the wells at various concentrations. Cells were harvested 48hrs after the addition of ligand and analyzed for luciferase and β-galactosidase activity. All data points represent the average of triplicate experiments normalized against β-galactosidase activity, standard deviations are represented using error bars.

In Silico Docking Studies

Modeling of ligand-receptor complexes were carried out using AutoDock Vina®. The crystal structure of human estrogen receptor β was obtained from the research collaborator for structural bioinformatics (RCSB) protein data bank (PDB). Ligands were created in ChemBioDraw Ultra 13.0 transferred to ChemBio3D Ultra where they were minimized using MMFF84 force field and saved as pdb files. Ligands were modified with Auto Dock Tools by adding Gasteiger charges to set the partial charge of each atom and hydrogens on polar atoms. Ligand- receptor docking was viewed using the Pymol molecular graphics system.

5.11 References

1. Kuiper GC, Enmark E, Pelto-Huikko M, et al. *Proc Natl Acad Sci USA* **1996,** 93(12), 5925-30.

2. Warner, M., Huang, Bo. Gustafsson, Jan-Ake. *Trends in Phamacological Sciences.* **2017**, 38

3. Tiziano Tuccinardi, Giulio Poli, Marco Dell'Agnello, Carlotta Granchi, Filippo Minutolo, and Adriano Martinelli. *J Enzyme Inhib Med Chem*, **2015**, 30(4), 662–670.

4. Chunyan Zhao, Karin Dahlman-Wright, Jan-Ake Gustafsson, Atlas Genet Cytogenet Oncol Haematol. **2009**; 13(2) 201-203.

5. Ascenzi, P.; Alessio, B.; Marino, M. *Molecular Aspects of Medicine*, **2006**, 27, 299-402.

6. Jia, M; Dahlman-Wright,K; Gustafsson, JA. *Endocrinology & Metabolism.* **2015**, 29, 557-568.

7. Lei CHEN etal. *Acta Pharmacologica Sinica.* **2014,** 35, 1333–1341.

8. John A. Katzenellenbogen *J. Med. Chem.* **2011**, 54, 5271–5282.

9. Anke Mueller-Fahrnow and Ursula Egner, *Current Opinion in Biotechnology*. **1999**, 10, 550–556.

10. De Angelis M, Stossi F, Carlson KE, Katzenellenbogen BS, Katzenellenbogen JA. *J Med Chem*, **2005**, 48,1132–1144.

11. Filippo Minutolo,1 Marco Macchia, Benita S. Katzenellenbogen, and John A. Katzenellenbogen, *Medicinal Research Reviews*, **2011**, 3, 364-442.

12. Hua-Jun Luo, Kun Zou, Nian-Yu Huang, Jun-Zhi Wang, Wei-Qiao Deng. *Med Chem Res*, **2013**, 22, 4468–4480.

13. McDevitt RE, Malamas MS, Manas ES, Unwalla RJ, Xu ZB, Millere CP, Harris HA. *Bioorg Med Chem Lett.* **2005**, 15, 3137–3142.

14. Williams,C; Edvardsson, K; Lewandowski, SA; Strom, A.*Oncogene*, **2008**; 27(7),1019-1032.

15. Chang, EC; Katzenellenbogen, BS. *Endocrinology*, **2006**, 147(10) 4831-42.

16. Julia Häring, Susanne Schüler, Claus Lattrich, Olaf Ortmann, Oliver Treeck, *Gynecologic Oncology*, 127 (2012) 673–676.

17. Ilaria Paternia,1, Carlotta Granchia,1, John A. Katzenellenbogenb, Filippo Minutoloa, *Steroids*, **2014**, 90, 13–29.

18. Gangadhara R Sareddy, et al. *Chin J Nat Med*, **2015**, 13(11), 801-807

19. Hsieh RW, Rajan SS, Sharma SK, Greene GL. *Steroids* **2008**, 73, 59–68.

20. Weiser MJ, Wu TJ, Handa RJ. *Endocrinology*, **2009**, 50, 1817–1825.

21. Meyers MJ, Sun J, Carlson KE, Marriner GA, Katzenellenbogen BS, Katzenellenbogen JA. *J Med Chem*, **2001**, 44, 4230–4251.

22. Sun J, Baudry J, Katzenellenbogen JA, Katzenellenbogen BS. *Mol. Endocrinol.* **2003**, 7, 247–258.

23. www.thermofisher.com/selectscreen

24. Richardson, T.; Norman, B.; Lugar, C.; Dodge, *J. Bioorganic & Medicinal Chemistry Letters.* **2007,** 17, 5563–5566.

25. Richardson, T.; Norman, B.; Lugar, C.; Dodge, *J. Bioorganic & Medicinal Chemistry Letters.* **2007**, 17, 5082–5085.

26. Richardson, T.; Norman, B.; Lugar, C.; Dodge, *J. Bioorganic & Medicinal Chemistry Letters* **2007**, 17, 4824–4828.

27. Richardson, T.; Norman, B.; Lugar, C.; Dodge, *J. Bioorganic & Medicinal Chemistry Letters, **2007***, 17, 3570–3574.

28. Filippo Minutolo,1 Marco Macchia, Benita S. Katzenellenbogen, and John A. Katzenellenbogen, *Medicinal Research Reviews*, **2011**, 3, 364-442.

29. M. Paz Otero, Alicia Torrado, Yolanda Pazos, Fredy Sussman, Angel R. de Lera. *J. Org. Chem.* **2002**, 67, 5876-5882.

30. Hinrich Gronemeyer, Jan-Åke Gustafsson, Vincent Laudet, *Nature Reviews.***2004**, 3.

31. Rebecca Skerrett a, Tarja Malma,b, Gary Landreth *Neurobiology of Disease*, **2014**, 72,104 – 116.

32. Bitoku Takahashi etal. *Journal of Medicinal Chemistry*, **2002**, 45,16.

33. Efrén Pérez, William Bourguet, Hinrich Gronemeyer, Angel R. de Lera. *Biochimica et Biophysica Acta, **2012***,18 (21), 57 –69.

34. A Szanto, V Narkar, Q Shen, IP Uray, PJA Davies, L Nagy. *Cell Death and Differentiation,* **2004,** 11, 126–S143.

35. Takemi Tanaka, Luigi M. De Luca. *Cancer Res*. **2009**, 69,12.

36. de Lera,A; William,B; Lucia,A; Hinrich, G. *Nature Reviews.***2007**, 6, 811.

37. S. Chandraratna. *J. Med. Chem.* **1995**, 38, 4764-4767.

38. Eric Finzi, Michael J. Blake, Paul Celano, John Skouge, Renuka Diwan. *American Journal of Pathology,* **1992**, 140, 6.

CHAPTER 6

IINVESTIGATION INTO FLUORESCENCE AND BINDING

6.1 Nuclear receptor probes

A major component of targeting nuclear receptors and successfully eliciting a desired physiological response is understanding its distribution within subcellular compartments. While nuclear receptors are known to localize in the cell's nucleus, receptors have been shown to exist in other areas such as the cytoplasm and mitochondria [1] then later translocating into the nucleus [2]. These findings open the door for the potential of targeting the nuclear receptor during different stages (pre or post translocation) however a clear understanding must be gathered on the nuclear receptors role outside of the cell's nucleus.

The vast implications of nuclear receptors in various diseases also strengthens the need for cellular tracking of them [3,4,5]. A better understanding of the receptors folding process in its native environment can provide insight into misfolding or constitutive turn on of receptors. As illustrated previously nuclear receptors undergo a cascade of events and a series of folding after ligand binding to turn on transcription all of which have a margin for error [6]. The ability to gain insight into the ligands role during conformation changes, the nuclear receptor response to the ligand and recruitment of co-activators can potentially aid in designing ligands with higher affinity.

Understanding of nuclear receptor movements within the cell can aid in better ligand design, as well as advancing the understanding of nuclear receptors role in various diseases. Several

methods currently exist to aid in providing insight into the subcellular movements of nuclear receptors.

6.1.1 Green Fluorescent Protein Fusion

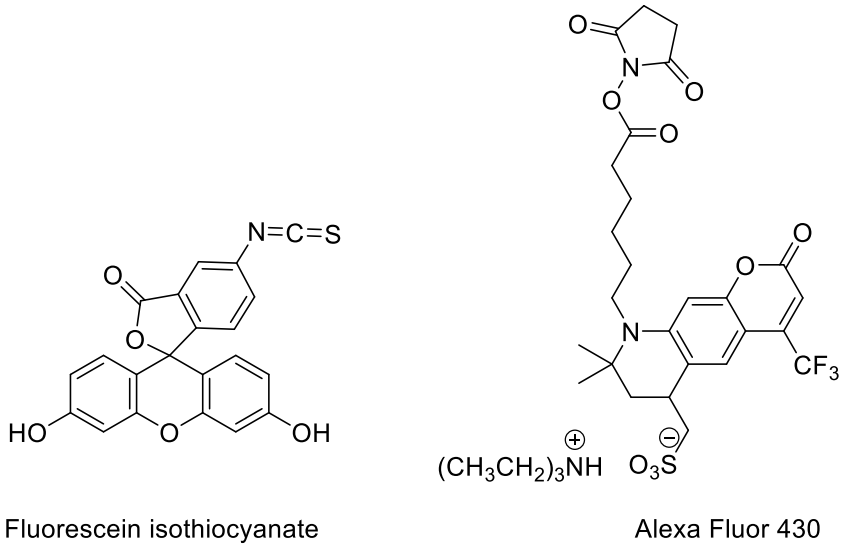
The discovery of GFP and its fluorescent variants has had a revolutionizing effect on most cellular visualization and trafficking applications. GFP fusion proteins have emerged as the most common and successful biotechnological application of GFP. Fluorescent fusion proteins have shown to boast several advantages over classical visualization probes such as GFP's lack of toxicity (in most cases), lack of requirement for co-factors or substrates and lastly the large diversity of fusion probes that can be generated. A detailed list of GFP fusion proteins have been published by Tsien [7].

Among the number of protein systems studied with GFP fusion are the nuclear receptors, of which several fusions have been made for the exploration of receptor mobility [8]. The ability to use GFP as a fluorescent tag successfully eliminated classical limitations and opened the door for advantageous *in vivo* subcellular trafficking. Several studies demonstrate the role and contribution that GFP has played in nuclear receptor mobility, compartmentalization, and ligand discovery with diverse nuclear receptors including GR, PR, and ER [9-11]. In general retention of transcriptional activity and ligand dependence is reported [12-15]. A disadvantage to fusion proteins specifically to the ligand binding domain is the potential interference with the ligand binding domain and the co-activator proteins essential for transcriptional regulation. Likewise, many probes are often bulky in size facilitating the previous disadvantage as well as a new one the decrease in affinity of the ligand to the receptor due to its steric bulk. Although subcellular visualization studies have been successfully conducted for several steroidal nuclear receptors the

possibility that tagging with GFP may disrupt some innate functions and localization of the natural receptor cannot be ruled out.

6.1.2 Classical Receptor Tracking

Classically receptor trafficking *in vivo* has been monitored *via* radioactive isotopes, enzyme-linked immunosorbent assay or ELISA and conjugate fluorescent dyes. While highly utilized, all the poses several limitations including radioactive hazards, fixation and incubation with antibodies and lastly size of the probe (Alexa Fluor® or DyLight Fluor®) [16].



Fluorescein isothiocyanate Alexa Fluor 430

Figure 6.1 Commonly used fluorescent tags. Commercially available as kits both Alexa Fluor and DayLight Fluor are fluorescent tags used for visualization, however both add bulk to their host.

Included in nuclear receptors that utilize traditional tagged ligands for monitoring are the estrogen receptors. Estradiol Glow [17], a fluorescent derivative of the endogenous ligand 17β-estradiol is commercially available. Structurally estradiol remains intact however the A-ring phenol has been modified for the inclusion of the tag. As shown in figure 6.2 modification involved deprotonation of the hydroxy moiety abolishing its ability to function as a hydrogen

bond donor. Formation of an O-C bond now renders functionality to hydrogen bond acceptor. In contrast tagged estradiol probes exist that maintain all functional groups of 17β estradiol however contains a large sterically bulky tag. As illustrated in our work and the work of others [18] though the estrogen receptor LBP can accommodate larger volumes ideal ligands do not occupy the entire pocket. The dilemma inclusion of sterically bulky tags adds is disruption to native folding behavior of the protein thus creating a demand for a novel class of probe that co-functions as fluorophore and ligand.

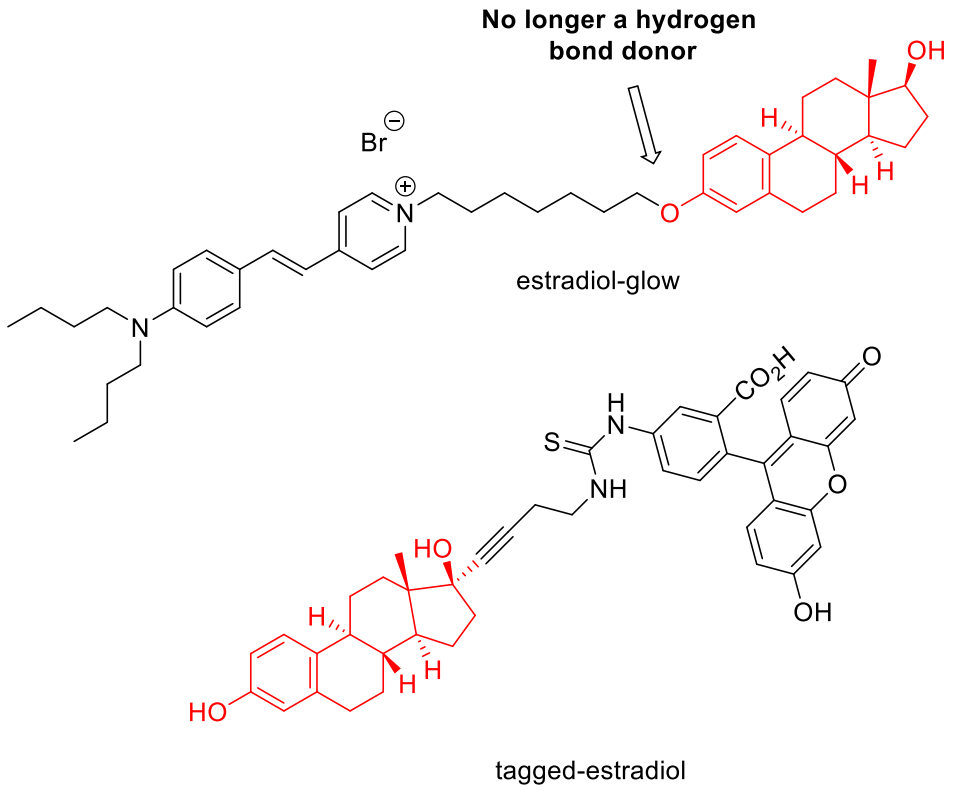


Figure 6.2 Classical Estradiol Modified Probes. Estradiol-glow requires altering estradiol (red) at the A-Ring phenol changing its functionality to a hydrogen bond acceptor as opposed to donor. The classical dye-tagged estradiol does not alter estradiol structure (red) however address steric bulk.

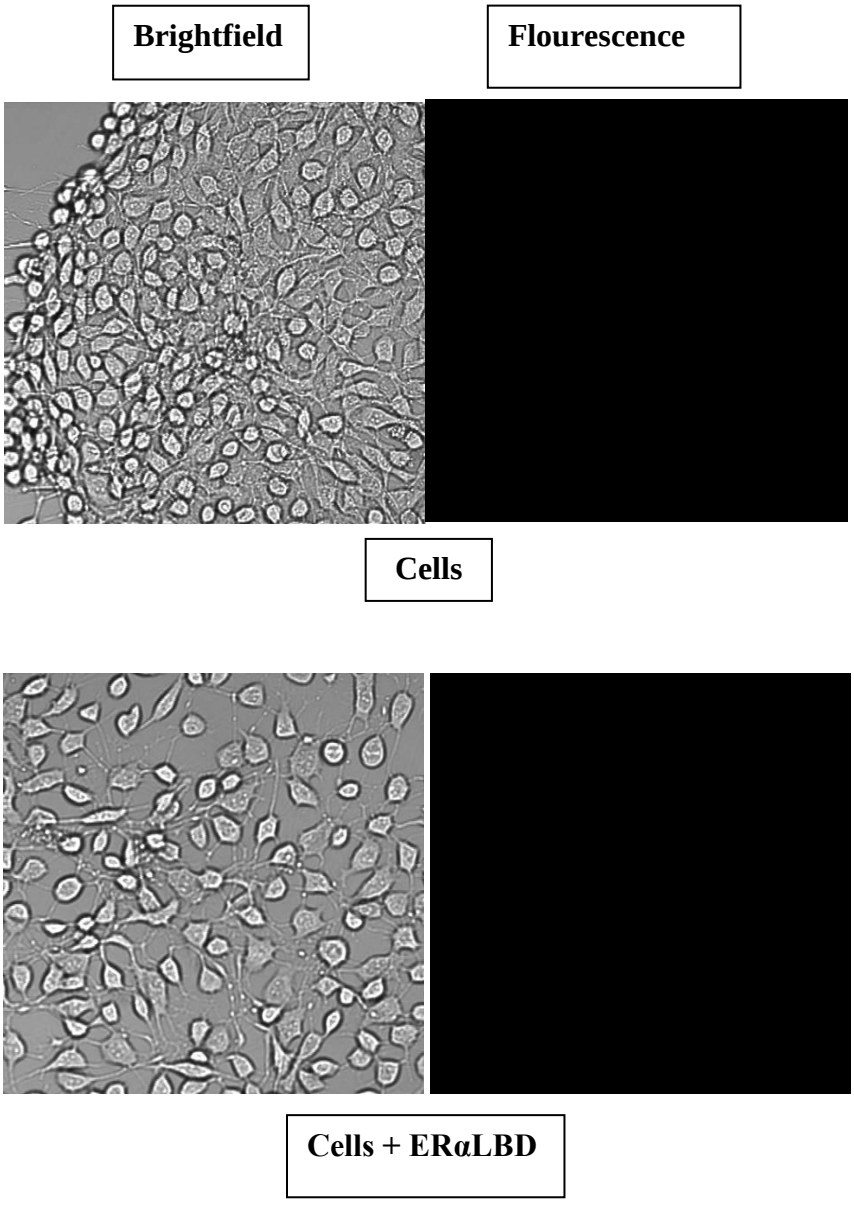
6.1.3 GFP inspired ligands as nuclear receptor probes

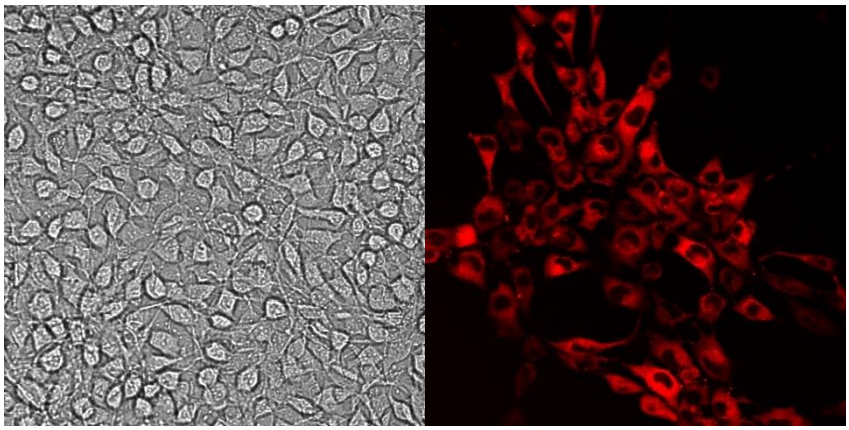
One of the rationales for using the GFP chromophore as the scaffold for receptor agonist was our hypothesis that the intrinsic ability to fluoresce could be re-activated upon binding thus tying binding to fluorescence. If the chromophore inspired ligands could invoke florescence within the nuclear receptor as a direct consequence of binding a new type of probe could be established. This class of fluorescent probe would boast several advantages over traditional forms of cellular tracking including the elimination of sterically bulky attachments and the handling of radioactive tags. The primary benefit of such a class of ligand is the ability to monitor the nuclear receptors movement without disruption of the receptors innate interaction with the ligand allowing for both in vivo and invitro cellular tracking.

6.2 GFP-Inspired ligands as nuclear receptor probes invivo

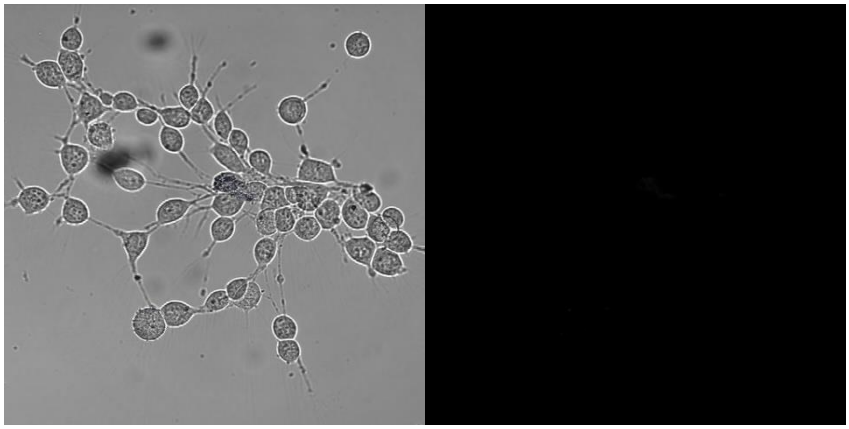
The hypothesis that the GFP-chromophore inspired ligands could co-function as a probe was assessed by transfecting NIH3T3 cells with plasmid pCMXwithERαLBD. After 8 hours of incubation in humidified air at 37^{o}C, cells were treated with select ERα active ligands and placed back in the incubator for 48 hours. At the end of 48 hours cells were removed washed with 1XPBS and viewed via confocal microscopy for fluorescence. To establish a background a set of cells were not transfected, a set of cells were transfected however not treated with a ligand and lastly a set of cells were transfected and treated with solely the vehicle. A set of cells were also treated with the ligands in the absence of ERαLBD to ensure that any observed fluorescence is attributed to just the presence of the ligand. As a positive control estradiol glow a tagged derivative of estradiol that fluoresces red was used. Lastly as a negative control ligand CW33 which has consistently shown no activation in ERα was used. Figure 6.3 below illustrates the

controls as the bright field image (left) and the image at the fluorescent image (right). Results are shown for the top activator CW32 and mid activator CW28.

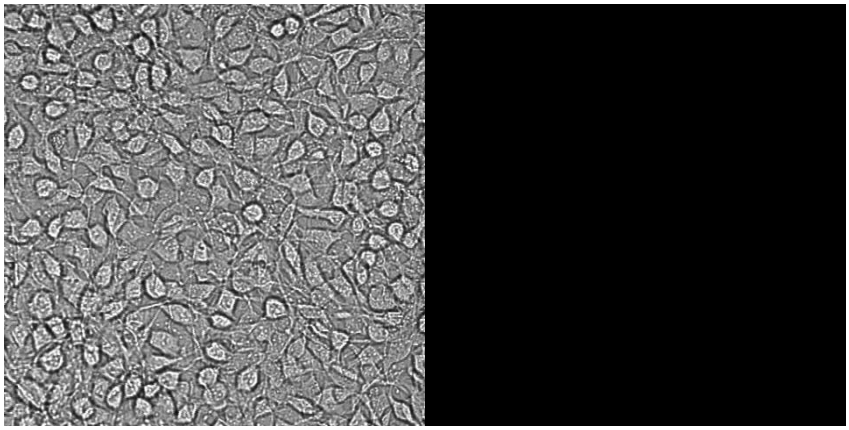




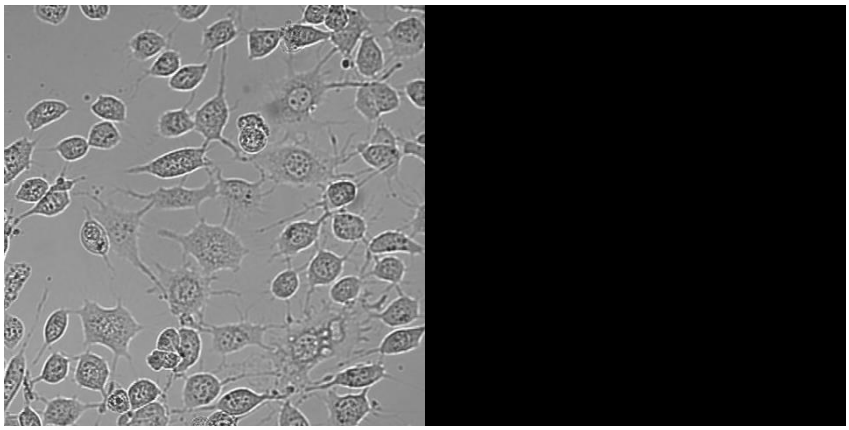
Cells + ERαLBD + Estradiol-Glow



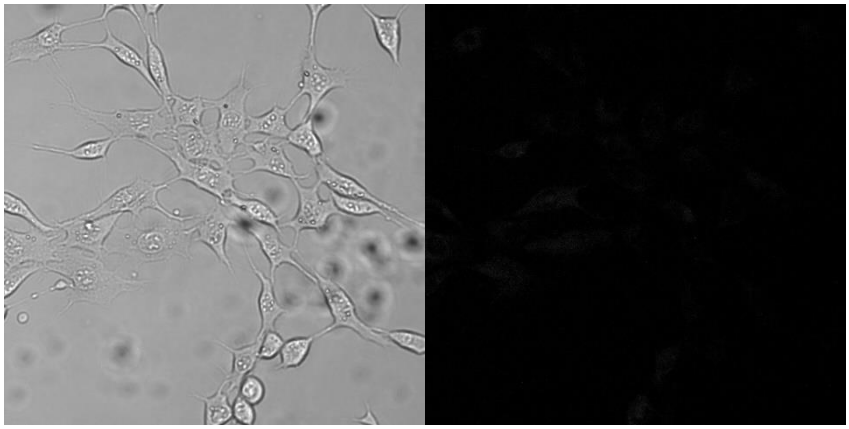
Cells + ERαLBD + CW33



Cells + CW32



Cells + CW32 +ERαLBD



Cells + CW28 +ERαLBD

Figure 6.3 Confocal images of GFP-chromophore inspired ERα ligands. The above images illustrate controls along with images for ligands CW32 and CW28. Images are displayed as brightfield (left) and fluorescent (right).

Analysis of the above images concluded that fluorescence was not observed in NIH3T3 cell line with the GFP inspired ERα top agonist ligand. Visualization studies were continued invitro via a collaborative effort with the Snell Lab.

6.3 GFP-Inspired ligands as nuclear receptor probes invitro

In a collaborative effort pioneered by the Snell Lab to demonstrate that estrogen like receptors exist in the *B.Manjavacas* [19] a species of Rotifera several arylmethyleneimidazolone ligands were submitted for a series of invitro testing. It is important to note that ligands submitted for testing pre-date this work however they have previously been shown to behave as agonist for ERα (unpublished results) [20]. Ligands submitted for invitro testing aided in the design of the new synthetic ligands generated in this work so while they are not the strongest activators they are a viable representation of GFP-inspired chromophores as ERα agonist. To ensure that submitted ligands were interacting with an estrogen like receptor in rotifers several tests were conducted including a reproduction test in which at least two ligands show increase in off spring generated. Once confirmed that the ligands were indeed binding with an estrogen like receptor the rotifers containing treated with the ligands were viewed using confocal microscopy. Images for several ligands in vitro are displayed below in Figure 6.4.

The results of figure 6.4 indicate that the GFP-inspired chromophore ligand can indeed be used both as a classical ligand and as a cellular tracking probe. Employed as a control estradiol-glow (red) certainly strengthens the hypothesis that estrogen like receptors are indeed present but also provides a template guide to locations in which this receptor could be located. Analysis of the confocal images indicate that in addition to fluorescence localization can be observed for certain ligands.

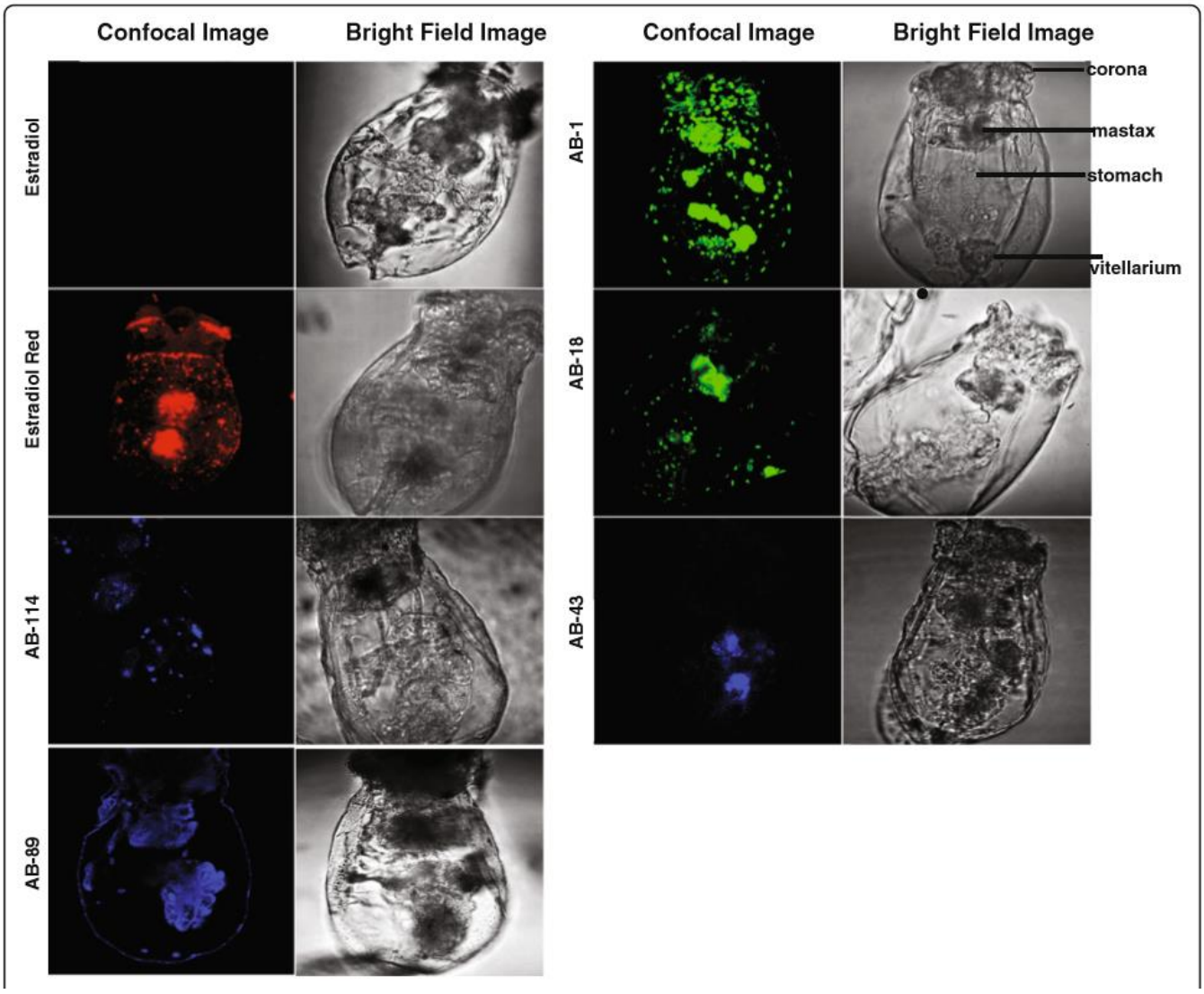


Figure 6.4 Synthetic fluorescent ligands bind selectively within B. manjavacas neonates. Bright field confocal and merged microscopic images of rotifers treated with ER binding small molecules. The compounds bind primarily to the reproductive tissues, the vitellarium, and the mastax. Estradiol and compound AB-89 exhibited greater binding in the vitellarium than in other tissues

6.4 SUMMARY

The investigation into the correlation between fluorescence and binding yielded interesting results. While in vivo studies showed no fluorescence the same did not hold true for invitro studies. Several ERα agonist that predate this dissertation demonstrated (unpublished results) fluorescence during invitro studies with the microinvertebrate rotifera. The GFP-inspired chromophore ligands not only demonstrated fluorescence but displayed localized fluorescence in specific systems within the rotifer. While the ligands that exhibited fluorescence do not activate with the same level of intensity as the ligands developed during this dissertation the mere fact that activation and fluoresce were housed in the same compound is proof of concept that the GFP-inspired chromophore ligand can indeed serve dual functionality. The combination of both results suggest that ligands can bind and activate however may not fluoresce likewise ligands may bind and fluorescence. The interesting thing to note is if higher binding and activation is associated with lower ability to produce fluorescence.

6.5 Materials and Methods

NIH3T3 cells were transfected with plasmid pCMXwithERαLBD using Polyfect. After 8 hours of incubation in humidified air at 37^{o}C, cells were treated with select ERα active ligands and placed back in the incubator for 48 hours. At the end of 48 hours cells were removed, aspirated and washed with 1XPBS and viewed via confocal microscopy for fluorescence.

6.6 References

1. Yang, S.H. *Proc Natl Acad Sci U S A*, **2004**. 101(12): p. 4130-5.
2. Hager, G.L., et al. *J Steroid Biochem Mol Biol*, **2000**. 74(5): p. 249-54.
3. Wagner, M., Zollner, G., Trauner, M. *Hepatology* **2011**;53: 1023-1034.
4. Schulman, Ira. *Advanced Drug Delivery Reviews* **2010**; 62: 1307-1315.
5. Skerrett, R., Malm, T., Landreth, G. *Neurobiology of Disease* **2014**; 72: 104-116.

6. Nicholes, C.; Liu, K.; Lin, C.Y. *Journal of Cellular Biochemistry*, **2013**, 114, 2203-2208.

7. Cubitt, A. B.; Heim, R.; Adams, S. R.; Boyd, A. E.; Gross, L. A.; Tsien, R. Y. *TIBS*, **1995**, *20*, 448-455.

8. Arun, K.H. S; Kaul, C.L.; Ramarao, P. *J. Pharmacol.Toxicol Methods*, **2005**, *51*, 1-23.

9. H. Htun, L.T. Holth, D. Walker, J.R. Davie, G.L. Hager, *Mol. Biol. Cell.* **1999**,10, 471–486.

10. Kawata, M; Nishi, M.; Matsuda, K.; Sakamoto, H.; Kaku, N.; Masugi-Tokita, M.; Fujikawa, K.; Hirahara-Wada, Y.; Takanami, K.; Mori, H. *Journal of Neuroendocrinology,* **2008***, 20*, 673–676.

11. Hager, G; Carol, S.; Lim, E; Baumann, C. *Journal of Steriod Biochemistry & Molecular Biology*, **2000**, *74*, 249-254.

12. V. Georget, J.M. Lobaccaro, B. Terouanne, P. Mangeat, J.C. Nicolas, C. Sultan. Mol. Cell *Endocrinol.* **1997**, 129, 17–26.

13. H. Htun, J. Barsony, I. Renyi, D.J. Gould, G.L. Hager. *Proc. Natl. Acad. Sci. USA.* **1996**, 93, 4845–4850.

14. C.S. Lim, C.T. Baumann, H. Htun, W. Xian, M. Irie, C.L. Smith, G.L. Hager. *Mol. Endocrinol.* **1999**, 13, 366– 375.

15. X. Zhu, J.A. Hanover, G.L. Hager, S.-Y. Cheng *J. Biol. Chem.* **1998**, 273, 27058–27063.

16. www.thermofisher.com

17. https://www.jenabioscience.com/probes-epigenetics/cell-labeling/fluorescent-hormones/pr-958-estradiol-glow

18. John A. Katzenellenbogen *J. Med. Chem.* **2011**, 54, 5271–5282.

19. Jones, B. *BMC Evolutionary Biology*. **2017**, 17, 65

20. Unpublished results: Duraj-Thatte, Anna. Flourescent GFP Chromophores as Potential Ligands for Various Nuclear Receptors. **2012**

CHAPTER 7

CONCLUSION AND FUTURE WORKS

7.1 Conclusions

Nuclear receptors role as transcription factors has caused them to be implicated in diseases ranging from cancer to diabetes [1]. Modulation of these transcriptional on/off switches is ligand mediated therefore providing a potential route to promote response from transcriptional machinery [2]. The potential to control key physiological responses at a subcellular level via the introduction of a ligand makes nuclear receptors of pharmaceutical significance [3]. Due to nuclear receptors wide distribution within the body [4] and commonality in structural motif [5], it is imperative when designing ligands that selectivity is exhibited by the ligand for not only separate families of nuclear receptors but also different forms within the same family termed isotypes.

Residing at the core of designing ligands that target nuclear receptors rest the interaction between the nuclear receptors ligand binding pocket and the ligand. Insight of the LBP has operated as the centerpiece for nuclear receptor binding research. Broadly, the LBP can be characterized as a cavity highlighted by nonpolar amino acid residues while possessing few polar residues. Divergent across receptors [6], varying in shape, size, volume and residues the disparity within the LBP provides the distinguishing element that allows differentiation between ligand features that are exchangeable from those that are indispensable.

Among the nuclear receptors that have gained considerable attention are the estrogen receptors. Notable for their roles in cancers [7] several ligands have been developed that target the specific isotype estrogen receptor α. Katzenellenbogen has significantly contributed to what is

understood about the estrogen receptor and ligand relationship, providing revelation into necessary functionality of compounds that bind within the ligand binding pocket [8]. Katzenellenbogen's devised pharmacophore for estrogen receptor activators highlights requirement of the phenol, flexibility within scaffold options (heterocyclic), and suggested substituents [9]. The ACD system pharmacophore Katzenellenbogen established has served as a blueprint for design of acceptable ligands with consideration to ligand size, volume, polarity, polar contacts.

This dissertation illustrates the synthesis of a small library of GFP chromophore inspired ligands as estrogen receptor α agonist. Chapter 4 highlights the use of Katzenellenbogen's pharmacophore as well as previous work conducted from our lab in this area to rationally design and evaluate additional GFP inspired ligands. Key findings include 10 ligands that serve as ERα agonist that were generated based on structural activity relationships from synthesis of multiple generations. Included within those 10 ligands are a super agonist CW32 that exhibits agonist behavior 235% that of 17β-estradiol the endogenous ligand for the receptor. Computational modeling using Auto Dock Vina supported rationale established for the successful agonistic behavior that careful consideration must be given to 1) hydrogen bonding network with the LBP, 2) the size and volume of the ligand and lastly 3) the molecular topography of the ligand. Consideration to these guidelines supported that high activating ligands give priority to participating in hydrogen bonding with either key histidine, glutamine, arginine residues or backbone residues as in the case of CW32. Contributing to activation are ligand volume and molecular topography with successful ligands of this class producing higher activation at volumes comparable or lower than that of estradiol. Lastly ligands with smaller nonpolar, branched moieties produced higher activation than those without.

Once successful generation of a series of ERα active ligands had been established, Chapter 5 illustrated the need for isotype and nuclear receptor selectivity. Key highlights of chapter 5 include conformation of preferential activation of several compounds including the top two activators for ERα. TFERT findings also included conformation that our computationally modeled designed ligands indeed bind in the same ligand binding pocket as that of the estradiol. While some ligands activated both isotypes with preference, ligand CW72 exclusively activated ERα. Lastly ERα active ligands exhibit no activation in the additional nuclear receptors test.

Understanding of binding, activation and receptor selectivity guided our use of the ligands as cellular probes. Originally inspired by the chromophore of the GFP Chapters 6 highlights the uses of these ligands as in vivo and in vitro probes for cellular tracking of nuclear receptors. Key highlights include the successful demonstration of this class of ligand to serve as both a binder/agonist and a fluorophore in vitro. Globally these findings support the hypothesis that GFP chromophore inspired ligands can co-function as both the stimuli and the probe. With the availability of other fluorescent protein chromophores and the broad range of available nuclear receptors this opens the door to a new class of ligand.

7.2 Future Works

While this dissertation covers a good breadth of work in the synthesis, biological evaluation and cellular application of GFP chromophore inspired ligands specifically for estrogen receptor α work remains to be completed. Future work can be separated into two categories 1) enhancement of binding through design and development 2) enhancement of cellular probe tracking applications.

To further the development of this class of ligand priority should be given to assessing the remainder of the ERα active ligands in TFERT assay against both isotypes. It is possible that other ligands exist in the set that are completely exclusive for ERα. Completion of this study will allow for a 3rd generation lead ligand that demonstrates complete selectivity for ERα currently that new lead would be CW72. While a drawback of CW72 is it low activation it is the only ligand thus far to show complete specificity for ERα. Consideration should be given on how to enhance binding and activation of CW72 while maintaining specificity. Concurrently due to CW32 high level of activation consideration should be given on understanding what makes CW72 specific and incorporating those features into CW32 will maintain super agonistic activation profile. Lastly with respect to ligand development the scope of this dissertation only covered agonist behavior however the same library of ligands should be screened for antagonist behavior. Antagonist, while they may elicit a different response, still bind and orient themselves in the ligand binding pocket providing more insight as to the development of optimal binders for this class of compound.

Development of cellular tracking and visualization should initiate with the revisiting of the current active set of ERα agonist. While initial in vivo visualization studies were conducted at various concentrations and different incubation times test, was limited to one cell line NIH3T3. The assay should be reconducted in an additional cell line specifically HEK293T. If accessible CW32 and CW72 should be submitted for in vitro testing in rotifera testing; CW32 due to its ability to function as a super agonist and CW72 due to its selectivity for solely one isotype of estrogen receptor.

7.3 References

1. Wagner, M.; Zollner, G.; Trauner, M. Nuclear receptors in liver disease. *Hepatology*, **2011**,53, 3, 1023-1034.

2. Sever. Richard, Glass. Christopher K. *Cold Spring Harb Perspect Biol.* **2013**, 5,167-09

3. Fraydoon Rastinejad.; Pengxiang Huang.; Vikas Chandra.; Sepideh Khorasanizadeh. *J Mol Endocrinol*, **2013**, 51.

4. Shabnam Farzaneh and Afshin Zarghi. *Sci Pharm*. **2016**, 84(3), 409-427

5. Katzenellenbogen, J.; Katzenellenbogen, B. *Chemistry & Biology* **1996**, 3, 7, 529-536.

6. Moras, D. *Current Opinion in Cell Biology*, **1998**, 10, 384-391.

7. Deroo, Bonnie.J.; Korach, Kenneth S. *Journal of Clinical Investigation*, **2006**, *116*, 561-570.

8. Anstead, Gregory; Carlson, Kathryn; Katzenellenbogen, John. *Steroids*, **1997**, 62, 268-303.

9. Katzenellenbogen, J. *Journal of Medicinal Chemistry*, **2011**, *54*, 5271-5282.

APPENDIX A

CHROMOPHORE PROPERTIES OF AMI LIGANDS

Ligands	Λ-max	Λ-excitation
Tetralin Core		
CW1	355	422
CW2	351	446
CW3	355	424
CW4	355	432
CW6	355	446
CW7	360	438
CW8	360	424
CW9	345	425
CW13	355	456
CW15	360	452
CW16	355	500
CW18	345	425
CW20	360	430
CW22	355	414
CW23	355	452
CW41	355	420
CW42	360	424
CW43	360	426
CW46	360	432
CW47	360	452
CW54	360	426
Napthyl Core		
CW48	368	435
CW49	363	452
CW50	369	431
CW51	370	429
CW57	349	449
CW59	369	429
CW60	368	455
N-propyl		
CW24	369	439

CW25	378	455
CW29	355	444
CW30	382	600
CW33	348	429
CW52	368	431
N-Isopropyl		
CW10		
CW27	380	450
CW28	372	445
CW31	353	440
CW32	382	445
CW72	357	510
N-phenol		
CW34	370	451
CW36	371	445
CW39	368	455
CW61	384	600
CW63	402	577
CW81	382	605
CW82	357	504
N-Cyclic		
CW35	374	451
CW38	370	457
CW68	399	569
CW69	401	571
CW76	383	600
CW77	384	600
N-Chiral		
CW40	370	447
CW55	374	449
CW84	370	451
CW85	370	595
CW86	370	455
CW90	396	567
CW91	370	483
CW92	375	451
CW93	357	446
CW95	357	450

CW96	383	450
CW97	383	595
CW98	378	600

APPENDIX B

SPECTRAL CHARACTERIZATION OF ACTIVE LIGANDS

CW28

^{1}H NMR δ: 1.46 (d, 6H, C(CH_3)$_2$), 2.41 (s, 3H, N=C-CH_3), 4.21-4.30 (m, 1H, N-CH), 6.82 (d, 2H, Ar-H), 7.02 (s, 1H, C=C-H), 7.99 (t, 2H, Ar-H). Calculated MS (ESI): 245.1 $[M+H]^+$

CW29

^{1}H NMR δ: 1.09 (t, 3H, C-CH_3), 1.75-1.83 (m, 2H, CH_2), 2.49 (s, 3H, N=C-CH_3), 3.70 (t, 2H, N-CH_2), 6.98 (d, 1H, Ar-H), 7.36 (s, 1H, Ar-H), 7.39 (t, 1H, Ar-H) 7.55 (d, 1H, Ar-H) 7.94 (s, 1H, C=C-H). MS (ESI): 245.1 $[M+H]^+$

CW30

^{1}H NMR δ: 1.85 (t, 3H, CH_3), 1.62-1.74 (m, 2H, CH_2), 2.38 (s, 3H, CH_3), 3.59 (t, 2H, N-CH_2), 6.81 (t, 1H, Ar-H), 6.92 (d, 1H, Ar-H), 7.15 (s, 1H, C=C-H), 7.31-7.39 (m, 2H, 2Ar-H). ^{13}C NMR: MS (ESI): 245.1 $[M+H]^+$

CW32

^{1}H NMR δ: 1.46 (d, 6H, (CH_3)$_2$), 2.39 (s, 3H, N=C-CH_3), 4.25 (m, 1H, N-CH), 6.82 (t,1H, Ar-H), 6.93 (d, 1H, Ar-H), 7.07 (s, 1H, C=C-H), 7.27-7.33 (m, 2H, 2Ar-H). MS (ESI): 245.1 $[M+H]^+$

CW34

^{1}H NMR δ:2.11 (s, 3H, N=C-CH_3), 6.82-6.85 (dd, 5H, Ar-H), 7.12 (dd, 2H, Ar-H), 8.08 (d, 2H, Ar-H, C=C-H). MS (ESI): 295.1 $[M+H]^+$

CW35

^{1}H NMR δ: 1.49-2.03 (m, 8H, $(CH_2)_4$), 2.47 (s, 3H, N=C-CH_3), 4.21-4.34 (m, 1H, N-CH), 6.79 (d, 3H. Ar-H), 8.02 (d, 2H, Ar-H, C=C-H). MS (ESI): 271.2 $[M+H]^+$

CW38

1.12-1.39 (m, 2H), 1.59-1.79 (m, 5H), 1.95-2.17 (m, 2H), 2.38 (s, 3H, N=C-CH_3), 3.61-3.77 (m, 1H, N-CH), 6.81 (d, 3H, Ar-H), 8.02 (d, 2H, Ar-H, C=C-H). MS (ESI): 285.2 $[M+H]^+$

CW39

^{1}H NMR δ:2.13 (s, 3H, N=C-CH_3), 3.79 (s, 3H, OCH_3), 6.80-6.83 (dd, 5H, Ar-H), 7.14 (dd, 2H, Ar-H), 8.11 (d, 2H, Ar-H, C=C-H). ^{13}C NMR: MS (ESI): 309.1 $[M+H]^+$

CW61

^{1}H NMR δ:2.13 (s, 3H, N=C-CH_3), 6.81 (d, 2H, Ar-H), 6.84 (d, 2H, Ar-H), 7.12 (d, 2H, Ar-H), 8.07 (d, 2H, Ar-H, C=C-H). MS (ESI): 295.2 $[M+H]^+$

CW63

^{1}H NMR δ:2.38 (s, 3H, N=C-CH_3), 6.12 (s, 1H, Ar-H), 6.21 (d, 1H, Ar-H), 6.82 (d, 2H, Ar-H), 7.12 (d, 2H, Ar-H), 7.18 (d, 1H, Ar-H), 8.23 (s, 1H, C=C-H). MS (ESI): 311.2 $[M+H]^+$

CW72

^{1}H NMR δ: 1.47 (d, 6H, $C(CH_3)_2$), 2.41 (s, 3H, N=C-CH_3), 4.23 (m, 1H, N-CH), 6.71(d, 1H, Ar), 6.83 (t, 1H, Ar-H), 7.11 (d, 1H, Ar-H), 8.01 (s, 1H, C=C-H). MS (ESI): 261.1 $[M+H]^+$